Accepted Manuscript

Health and Education Expansion

Jonathan James

PII: S0272-7757(15)00117-X
DOI: 10.1016/j.econedurev.2015.10.003
Reference: ECOEDU 1593

To appear in: *Economics of Education Review*

Received date: 7 September 2014
Revised date: 21 September 2015
Accepted date: 15 October 2015

Please cite this article as: Jonathan James, Health and Education Expansion, *Economics of Education Review* (2015), doi: 10.1016/j.econedurev.2015.10.003

Highlights

- Proportion of 18 year olds in full time education rose from around 17
- The expansion resulted in a rapid increase in education over the whole education distribution.
- Education led to a reduced BMI, waist circumference and weight.
- However, no effects are found for smoking, drinking or hypertension.

Health and Education Expansion*

Jonathan James

University of Bath

October 23, 2015

Abstract

In this paper I exploit a reform that expanded UK post-compulsory education during the 1980s and 1990s to examine the effect of education on health. The expansion resulted in a rapid increase in education over the whole education distribution. I find evidence that education had an effect in reducing body mass index, waist circumference and weight. For other health measures (self-reported general health, long term or limiting illnesses), blood pressure and health behaviours (smoking and drinking) there were small to no improvements. There is suggestive evidence that the mechanisms driving these results are improvements in labour market and social status.

JEL Classification: I20, I10, J10, J20
Keywords: Health, Education

*Correspondence address: University of Bath, Claverton Down, Bath, BA2 7AY. Email: j.james@bath.ac.uk. I would like to thank seminar participants at the European Association of Labour Economists conference 2013 and the Applied Economics of Education Workshop 2013 as well as two anonymous referees.

1. Introduction

Across developed countries there have been large increases in post-compulsory schooling during the last few decades. The UK saw a particularly rapid rise over a short period of time. From the late 1980s to the early 1990s the proportion of 18 year olds in full time education rose from around 17% in 1985 to over 35% in the late 1990s. There has been increasing interest in the effects of education beyond the effect on wages from within both academia and among policy makers. This paper examines this rapid post-compulsory education expansion in the UK to examine the effects on health.

Many recent studies have used quasi-experimental evidence, in particular using schooling laws in an instrumental variables setting or using regression discontinuity (RD) designs, (Lochner 2011). For example, Clark and Royer (2013) examine the effect of two compulsory schooling law changes. In the UK the minimum school leaving age was increased by one year to fifteen in 1947 and to sixteen in 1973. They investigate the effect of these changes on mortality, other health outcomes, and behaviours. They find insignificant effects of these law changes. Lleras-Muney (2005) used changes in compulsory schooling laws and child labour laws in the US exploiting the variation in timing and geographical implementation of these laws. She finds additional schooling leads to reductions in mortality although this is not robust to the inclusion of state-specific trends, (Mazumder 2008). Albouy and Lequien (2008) use a RD strategy similar to Clark and Royer (2013) and also do not find any effects on mortality using French data. In contrast, van Kippersluis et al. (2011) do find effects on mortality by exploiting a change in the Dutch compulsory schooling laws. Their large sample size and strong effect of the law change on schooling means that they are able to detect small effects which other studies may not be able to. Arendt (2005) uses a Danish reform and panel data to find some evidence of an effect on self-reported health and smoking.

Turning to other measures of health Oreopolous (2006), and Silles (2009) also examine compulsory law changes in the UK and find reductions in the probability of self-reported poor health. This is in contrast to Clark and Royer (2013) who use month of birth, rather than year of birth, and account for cohort trends. Powdthavee (2010) examines an important biomarker – blood pressure. He finds that education reduces hypertension for men. Jurges et al. (2012) estimate no significant effect of the 1947 and 1973 law changes on two other biomarkers that are relevant for heart disease, C-reactive protein and blood fibrinogen. Braakmann (2011) exploited a discontinuity in the month a student could leave school which resulted in an increase in the probability to obtain a qualification but found no impact on health outcomes or behaviour. Arendt (2008) exploits a Danish schooling reform to examine the effect on hospitalisations. Whereas these studies rely on the exogenous increase in education that almost exclusively had an effect at the bottom of the education distribution, this paper examines a reform that lead to an increase throughout the education distribution.

There are four main reasons why the expansion considered in this paper can provide valuable new evidence on the relationship between education and health. First, the reform was very large and resulted in a doubling in post-compulsory schooling participation. Second, the reform lead to changes throughout the distribution. There were changes at both the upper and lower ends of the education spectrum. Therefore, given this unusual distributional impact we may expect that this reform to have a different impact on health. Whereas the compulsory schooling law changes compel just those at the bottom of the distribution to stay on, it has been shown by Bladen and Machin (2007), that this particular education expansion mainly, although not exclusively, benefited those from better-off backgrounds. This reform is an interesting natural experiment as it results in estimating a different local average treatment effect (LATE) than in, for example, Clark and Royer (2013).

Third, the timing of the reform is relatively recent. Therefore, many of the health dangers of smoking, and drinking were becoming, if not already were, well known at the time of the reforms and therefore it may be the case that education has a less important role to play. Therefore, examining this reform may give a better insight into the relationship between health and education today. Finally, there is a policy relevant aspect to this reform. Expansion of higher education is a policy target of many countries. If a spill-over from this is improvements in health then this should be taken into account when forming higher education policy.

The results are somewhat mixed. There is little or no impact on self-reported health measures (general health and long standing illness), although there is a reduction in limiting illness overall. There is no effect on 'bad' health behaviours such as smoking and heavy drinking. However, I do find there is an impact of education in reducing body size shown by BMI and waist size. These results are robust to a range to different specifications. Mechanisms that could be driving these findings are also examined. In addition to the income effects that the reform lead to as shown by Devereux and Fan (2011) and Machin et al. (2012) I also examine social class and economic activity. While there were improvements in social class, with more moving into the professional, and managerial classes, there was not an equivalent change in economic activity. Other mechanisms such as peer effects due to partners who were affected by the reforms are examined but there is no evidence to suggest that they have an impact.

The paper proceeds as follows: section 2 explains the features of the expansion in education. Section 3 describes the data and descriptive statistics including a graphical analysis. Section 4 sets out the empirical strategy. Section 5 presents the results. Section 6 concludes.

2. Institution Setting: Education Expansion in the UK

In the UK the proportion of 18 year olds in full time education has rapidly expanded since the mid-1980s. Figure 1 shows the rapid increase in participation over the period, represented by a significant step change. The rise in full time participation increased from 17% in 1985 to 40% in 1995. The effect was not just on the lower part of the distribution. Figure 1 also shows that the rise occurred for both further education i.e. post compulsory schooling (the law in place at the time prohibited leaving school before 16) and higher education. For both measures there was over a doubling in participation over the period. In section 3 I will examine the degree of the changes in more detail using the Health Survey of England.

There were two main causes for the rapid rise in education over this period. First, there was a significant change to the high school exam system. Second, there were significant changes to the supply in higher education. The main high school exams were replaced, Blanden and Machin (2004). The General Certificate of Secondary Education (GCSE) combined the O-level (General Certificate of Education (GCE)), a higher tier exam, and the Certificate of Secondary of Education (CSE), a lower tier exam. These changes led to an improvement in results, which may have encouraged people to stay on into further and higher education. There were two main reasons why this exam change may have led to an improvement in attainment. First, a cap on the number of people who could receive a specific grade was removed. Therefore, more could achieve grades A to C which are considered to be passing grades. This move from norm-reference exams, which placed emphasis on relative performance, to criterion-referenced assessment meant that it was possible for everyone to get the top grades, Blanden et al. (2003). Second, there was a move away from the assessment being based just on exam performance to include a sizeable element of coursework.

Gray et al. (1993) show that the most important determinant in predicting post-16 schooling were the qualifications they had received. Gray et al. (1993) also provide evidence that there were big jumps in attainment. Using the Youth Cohort Study (YCS) they find that 30% obtained 4 or more high grade passes in 1986 (pre-GCSE), this increased to 40% in 1988 the first year of the GCSE. There was an increase at almost every level. Therefore someone born in 1972 and after who had the same ability and other similar characteristics (such as a similar discount rate) that someone born before 1972 would have had a greater opportunity to stay on in education due to the change in the examination system as they would have achieved the grades that would have allowed them to go on to further study. Changes to the structure of the economy, moving away from manufacturing and into services and the perceived increases in returns to education was also another significant driver of this increase in education demand, Blanden and Machin (2004), Kogan and Hanney (2000), Devereux and Fan (2011).

In addition to the high school exam system there were changes to the supply of university places. There were two key features that lead to the rise, (Walker and Zhu 2008). A relaxation in the limits of student places but also the government grant per student to universities fell and as a result this increased the incentives for the universities to enrol more students. The Further and Higher Education Act 1992 also lead to an expansion of university education as many polytechnic institutions became universities. The difference between the two institutions being that universities could award their own degrees.

The effects that are examined in this paper are therefore a combination of both policy changes. They do interact with each other and did not have independent effect on the education distribution. Alone, the exam changes would have had an impact at the bottom of the education distribution. However, with the expansion in higher education occuring in conjuction with the changes to exam system high school students then in a position to take advantage by staying in school longer and then moving into higher education. Devereux and Fan (2008) point out that the improvement in grades would have led to students believing they were good enough to go on to higher education. In section 5D we examine the different margins over which the reform would have had an effect on education and then subsequent health.

3. Data and Descriptive Statistics

I use the Health Survey of England from 1991-2012[1]. This is an annual cross-section survey that combines questionnaire with information collected more directly on health by a nurse survey. There are a number of key features of this data. First, it contains information on years of education and qualifications. Second, the health survey collects data on a wide range of health outcomes and behaviours. I consider three measures of self-reported health that give an overall picture of the state of the individual's health: general health, having a long standing illness, and having a limiting illness. In addition to broad health assessments I also investigate less subjective measures of health, these include BMI (body mass index) calculated from height and weight data collected by trained nurses. An important biomarker is also considered - high blood pressure. One issue is that the age of individuals sampled are probably too young for certain health conditions to materialise as I restrict the sample to those aged between 23 and 34 in order to keep a balance between treated and untreated individuals of ages in the sample. Therefore, I examine a range of self-reported health behaviours, such as measures of smoking and drinking. The cohorts included were those born between 1962 and 1980[2] to capture the cohorts before, during

[1]It is not possible to use all years of the Health Survey in all the analysis as some variables are not available in each year. All survey years are used except when the following variables are used: limiting illness (1996–2011), drank 7 days of the previous week, and drank over limits on heaviest drinking day (1998–2012)

[2]Where possible I assign individuals to the school cohort. Month or birth is available in the HSE in 1991, 1992, 1998, 1999, 2000, 2001 and 2002. For individuals born in September, October, November or

and after the expansion. Descriptive statistics of the age and education variables are presented in Table 1 and the health variables are presented in Table A1.

Next, I turn to a graphical analysis of how education has changed by cohort. Figure 2 presents the change in a number of education measures by cohort. I examine four different measures of education. The top left panel of Figure 2 shows the mean age left school for all individuals and broken down by men and women. For the cohort born in 1969 the average leaving age is around 17, throughout the period of the expansion this raises to 17.7. The expansion is greatest for the cohorts between 1972 and 1975 inclusive, as represented by the vertical lines on the graph.

Figure 2 also presents the proportion staying on after aged 16 in the top right panel. There is a similar pattern to the age left school. The increase is around 20 percentage points for both sexes, from around 50% for the 1969 cohort. After the 1975 cohort there is a plateau, with only the final 1980 cohort dropping, due to composition effects of the survey. The bottom two panels present cohort level changes in the proportion holding A levels or above and a obtaining a degree. There is a positive trend in both measures. The upward trend before the expansion period was greater for women, and both sexes experienced an increase above the trend during the expansion, and then reaching a plateau. There is a larger impact on those holding a degree, shown in the bottom right panel. The upward trend is greater for women before the expansion. During the expansion, the proportion with a degree rose from around 23% to 35% for men, and from just over 20% to 30% for women. Therefore, these figures show that there was a increase in education not just at the bottom of the distribution but also an increase higher up resulting in more people achieving degree level qualifications.

To further explore the cohort variation seen in Figure 2 I examine these trends in a different way. I assume that polynomial in age up to cubic to capture the secular trend in education. The residuals from these trends are presented in Figure A1. A similar pattern is seen in these graphs the positive deviation from the secular trend is mostly seen in the later cohorts for age left full-time education and leaving school post-16, for having A-levels and above, and having a degree, this positive deviation happens also for the earlier cohorts.

In addition to the changes in education by cohort I also present the analogous graphs for the health variables[3]. As in Figure A1 in order to take into account age effects which could well be the dominant effect I control for a polynomial in age up to a cubic[4]. Figure 3 plots three self-reported health measures. Being in good or very good health, having a long standing illness, and having a limiting illness. There are negative deviations from

December they are assigned to the following year.

[3]Descriptive statistics of the health outcome variables and health behaviours are presented in Table A2A and A2B.

[4]I have also tried specifications that included age on its own, and age and its square, qualitatively the graphs looks very similar.

the secular trend during education expansion period for being in very or good health. For having a long standing illness there appears to be a positve deviation at the beginning of the expansion period and then slightly negative ones in the latter cohorts, particularly for males. This pattern is similar to that seen for having a limiting illness although the deviation is for both male and females.

Figure 4 shows the body measurements. The top panels display the residuals for body mass index, overweight and obese. For all three variables there is a decline during the expansion period is not in line with the general trend. This appears to be more the case for the later cohorts of the expansion which is what one would expect given the gains in education were seen more for the later cohorts. The post-expansion cohorts are characterised by negative deviations until the last two cohorts. For waist and waist to hip ratio, the data is much noisier and the pattern is less clear in the expansion period, the immediate post cohorts after the expansion do display a negative deviation from the secular trend.

Figure 5 presents two blood pressure measures. The left panel shows the residuals for hypertension and the right panel shows the same with additionally controlling for drugs that influence blood pressure. There is some evidence of a positive deviation from the secular trend in the right hand panel particularly for the earlier cohorts. Figure 6 examines smoking and drinking behaviours. There is little evidence of a deviation during or after the expansion for the post expansion cohorts for either measures of smoking. For three measures of alcohol use there is also little evidence of a deviation from the secular trend.

4. Empirical Strategy

I begin by presenting the first stage showing the relationship between the education expansion cohorts and educational achievements:

$$Ed_{ic} = \alpha + \sum_{c=72}^{75} \beta_c \text{Cohort}_c + \delta After_c + f(Age_{ic}) + g(Cohort_c) + \epsilon_{ic} \tag{1}$$

Where subscript i represents the individual in cohort c. Ed is a measure of education. As in Machin et. al (2011) I parameterise the age polynomial with a cubic and cohort polynomial with a quadratic[5] The expansion in education is captured by the individual cohort dummies from 1972-1975, the coefficients give the difference in education and

[5]While this is not strictly an RDD in has a similar flavour, Gelman and Imbens (2014) point out that high order polynomials should not be used in RDDs. In any case, I test the robustness of the 2SLS estimates to the specification of both the "running variable" (birth cohort) and how age is put into the model. In figures 7–10, columns 9, 10, 11 include age quartic, year of birth cubic, and then both age quartic and cubic in year of birth. The results appear robust to the specification in this regard.

health relative to the pre-expansion cohorts. Furthermore, these dummies capture the deviation of the cohort trend given I include a polynomial in the cohort. $After_c$ is a dummy representing the cohorts after the expansion to capture the leveling off that was seen in the set of education graphs.

The identification strategy relies on examining cohort-level changes in education and a number of health outcomes. A nice feature of the reform in question is that it affected the entire education distribution. Unfortunately, however, this means that it is not feasible to use a portion of the cohort not affected as a control group, as in Etile and Jones (2011). Additionally, during this period of expansion Scotland also experienced expansion in the higher education sector so does not make a feasible control group. The strategy employed therefore rests in identifying changes in the cohort trend that cannot be captured using a low-order cohort polynomial. Therefore there maybe be underlying differences from cohort to cohort in health behviours, however, there is no reason to think that the other factors that influence health do not change smoothly, and would therefore be captured by the cohort trends.

In the second stage I use 2SLS to estimate the effect of education on health. Where education is determined by the first-stage equation (1) and a set of cohort dummies for those in the expansion period are used as excluded instruments as well as a post-expansion dummy.

$$H_{ic} = \lambda + \theta \widehat{Ed_{ic}} + h(Age_{ic}) + k(Cohort_c) + \nu_{ict} \tag{2}$$

The coefficient of interest is the effect of education on health, θ. The interpretation of this estimate, under the assumption of monotonicity, is a local average treatment effect (LATE), i.e. the estimated effect is for those who obtained more education as result of the expansion. In contrast to changes to compulsory schooling which exclusively affect the bottom of the distribution, this reform is for a broader part of the population, however, the interpretation remains the same in that the effect is for the compliers of the reform.

The validity of the instruments in this case rest on the assumption that the education expansion cohorts significantly explain the variation in education without being correlated with unobservable characteristics that are correlated with education and health such as family background, risk aversion, or time preference. Table 1 explicitly tests the first requirement. The second requirement needs that the expansion was not explicitly aimed at improving health or implemented as a reaction to poor health. There is no evidence that this is the case, for example the further and higher education act does not explicitly mention health outcomes as a reason for the changes either directly or indirectly. [6]

[6]Further and Higher Education Act 1992: www.legislation.gov.uk/ukpga/1992/13/pdfs/ukpga_19920013_en.pdf

The key identifying assumption in using the set of cohort dummies as instruments for education is that the conditional expectation of the health outcomes with repsect to the birth cohort is that in the absence the education expansion then the changes could be explained by a low-order cohort polynomial. Therefore, one way in which we can indirectly test this assumption is by examining the effect of the instruments (cohort dummies) on pre-determined characteristics. Testing these effects examines whether there are other factors changing over time that could explain the results. Specifically, Table A3 tests for deviations in the cohort trends for the proportion of stillborn births, infant mortality under 1 year and under 4 weeks, and births outside marriage. The results do not show any effect of the cohort dummies.

As discussed in section 2 what is being picked up is the effect of two policy changes. A change in the exam system and changes to the supply of higher education provision making it more accessible. The exam changes meant that someone born in 1972 or after who had the same ability and other similar characteristics (such as a similar discount rate) as someone born before 1972 would have had a greater opportunity to stay on in education due to the change in the examination system. This allowed students to remain in education longer who otherwise would have been allowed to. One concern could be that what is being captured is just a general cohort trend increase in education. To this end in section 5F I present a number of falsification tests. In particular, I replace the cohort dummies with a sets of cohort dummies prior to the reform. To further check the robustness in order to account for the possibility that other factors that could have happened at the same time as the expansion section 5F also presents the estimates when health spending (as a proportion of GDP and when the individual was 16) is controlled for. This attempts to capture not only the potential health effect of the additional spending but also general sentiment of the government that would be with regards to promoting health. In addition, there is a large literature on business cycles and health, therefore, I control for the unemployment rate at the time of the survey. These are presented in the last bars (no. 27) of figures A2-A5. These checks confirm our baseline results suggesting that effect of education is being captured.

5. Results

A. The Effect of Education Expansion on Education Outcomes

Table 1 shows the reduced form estimates for each of the education measures considered. The expansion in education is represented by the dummies for each of the 1972 to 1975 inclusive cohorts, as seen in the previous figures these were the years where the expansion was at its greatest. As described above I also include a dummy representing the post expansion cohort, therefore, the estimates I find are relative to the cohort who experienced the pre-GCSE exam system.

For all the measures of education there is an increase for each of the cohorts with each subsequent cohort being greater than the previous. For age left full time education the significant step change monotonically increases from 0.172 to 0.534 with the post expansion dummy being 0.546. The pattern is similar for staying on past 16, going from 0.0394 to 0.170 with a plateau at 0.184. There are improvements in qualifications, both achieving an A-level and above and obtaining a degree also display a similar pattern, albeit the F-statistic for the joint test of the 1972 to 1975 cohorts are not as large as compared to age left school or leaving education after 16 after. For the A-level and above measure the F-test is significant but not above 10. The F-tests for age left education, and post 16 are both above 10.

B. Health Outcomes

Self Reported Health

Table 2 presents reduced form, OLS and 2SLS estimates for three self-reported measures of health. The dependent variables are binary with a 1 if self-reported health is good and 0 otherwise, 1 if the respondent has a limiting illness 0 otherwise, and 1 if they report having a long standing illness. In the top panel I report the reduced form effects for the cohort dummies. For having a good general health and having a long standing illness there are no significant results. There is an impact on reporting a limiting illness, the estimates are not monotonically declining in line with the changes in education. The p-value of the F-test for all measures are all above 0.1.

OLS estimates are positive and significant suggesting that those with greater education are more likely to report better general health, leaving school one year later is associated with a probability of very good or good general health by 3 percentage points. A similar pattern, albeit with opposite sign, is found for having a long standing illness and having a limiting illness. As mentioned above allowing for the assumption that the education expansion only affected health through the education then we can estimate the health returns to education using 2SLS. The 2SLS estimates for general health and having a long standing illness are all imprecisely estimated. The IV estimate of education on long standing illness is around 50% larger that the OLS estimate, however the standard errors also become larger. The results suggest an increase of education by one year reduces the probability of a limiting illness by 5.7 percentage points. The magnitude of the effects is in line with Silles (2009) and Oreopolous (2007).

Body Size

Next I turn to directly measured health outcomes. Table 3 presents the results for BMI, overweight (BMI>25), obesity (BMI>30), waist measurement (cm) and waist to hip ratio. Each of the p-values of the F-tests testing the joint significance of the cohort dummies and post expansion indicator were all below 0.05 except for overweight which is below 0.1.

These results suggest there is a significant difference in body size for these cohorts relative to the pre-expansion cohorts. OLS estimates point to a negative and significant impact of education, such that leaving school a year later leads to a fall by -0.289 (kg/m^2), however the 2SLS estimate is larger at -1.48 (kg/m2), as a proportion of the mean BMI for the pre-expansion cohorts the 2SLS esimates represents a 5.4% fall in BMI. I reject the null that the OLS and 2SLS estimates are the same. In Table 3 all of the 2SLS estimates are larger than the OLS.

There are a number of reasons why this might be the case. Measurement error of the education variable can cause the OLS estimates to be biased towards zero. Below I discuss the results for smoking and drinking which suggest that measurement error is not driving the results I find. Alternatively, the marginal benefit of schooling for individuals whose education has been changed by the education expansion reforms may be larger than the average benefit for the population. The 2SLS estimates are a local average treatment effect (LATE), so it is the impact on those who took an extra year of education for those who would not have taken it if the expansion had not taken place. Those affected by changes to the exam system and higher education may have high marginal returns to schooling in terms of the adoption of healthier behaviours.

A one year increase in schooling reduces the probability of being overweight by 11.8 percentage points, off the pre-expansion mean this represents around a 15% decline. For obesity (BMI>30), as with the overweight estimates the OLS estimates imply that more education is associated with a lower probability of being obese. An alternative measure of body size I examine is waist and waist to hip ratio. There is a reduction in both waist and waist to hip ratio seen in the OLS estimates. The 2SLS estimate of the effect of education on waist to hip ratio is also negative but imprecise, the waist measure implies a reduction of around 2.5 cm as a result of staying in school an additional year. This is in line with the results found on body mass. In Tables 6, A6 and A7 I examine further the make up of this result by examining height and weight.

Blood Pressure

An additional objective health measure examined is blood pressure, and in turn hypertension. Table 4 presents the estimates of this analysis. The first column does not control for taking medication that could have an effect on hypertension, whereas the second specification does. Both the 2SLS estimates suggest an additional year of education reduces the incidence of hypertension but this is not statistically significant. One explanation for this, other than there being no actual effect, could be the sample is too young to pick up any effects on hypertension and the incident rates for this age of the population too low. The oldest individual in the sample is 34 which might be too young for this issue to have an impact. The reduced form effects also suggest that there was not a change in the measure of hypertension for these affected cohorts.

C. Health Behaviours

Smoking, and Drinking

Two measures of smoking are shown in Table 5. OLS estimates suggest a positive relationship between education and never smoking and negative in relation to being a current smoker. The 2SLS estimates are both negative but not statistically significant. By the time of the natural experiment considered here occurred the dangers of smoking were well known, indeed most of the subjects in this study were not even born when the 1964 U.S. surgeon general's report on smoking and health was published, and so the causal channel for formal education to play a role is therefore potentially diminished, hence it is maybe unsurprising that effects on smoking are not found. I also do not find any consistent reduced form effects.

Table 5 also examines three measures of alcohol consumption: being a current drinker, whether on the heaviest drinking day the individual drank above government recommended limits, and whether they drank 7 days a week in the last week. The OLS estimates for being a current drinker are positive, however, education is negatively associated with binge drinking (consuming 8 or more units). None of the 2SLS are precisely estimated and I do not find any effects from the reduced form analysis.

While the 2SLS estimates on body size and hypertension were larger than the OLS. That is not the case when I examine smoking and drinking. One interpretation is that educated individuals have unobserved characteristics that also make them more healthy or more likely to engage in healthy behaviours. Implementing instrumental variables controls for these characteristics and reduces the estimated effect of education on smoking but does not eliminate it. As discussed above the 2SLS estimates of the effects of education on body size are larger than the OLS estimates. If these results were driven only by measurement error of the individuals education, then one might expect IV to be systematically greater than OLS for all the dependent variables.

D. Heterogeneity

Table A4 presents 2SLS estimates separately for males and females for health outcomes (self reported health, blood pressure, and body size)[7] and Table A5 presents the same for health behaviours. The 2SLS estimate of the effect of education on limiting illness are both negative for men and women albeit only precisely estimated for the latter. As we have in the pooled sample there are not any significant effects for the other measures of self-reported health.

The effect on body size are prevalent for both men and women. These effects are in contrast to Webbink et al. (2010) who find no effects on body size for women using twin data, however, they are in line with Brunello et al. (2013) and Grabner (2009). For

[7]These results should be treated with caution as the first stage in a number of cases is below 10.

men the effect on body size also results in reductions of obesity but this is imprecisely estimated for women. The reductions in waist size also seem to be driven by males. I find a reduction by 3.7 cm as a result of staying on in school for an additional year. The 1.8 cm reduction in waist size for females is not precisely estimated.

Turning to Table A5, while there appears some heterogeneity in the results, for example, I find opposite signs on the effect of education on smoking behaviour for both males and females, none of the results are significant.

In addition to heterogeneity of the effect with respect to gender, I also consider different definitions of education. Figure 7 shows the 2SLS estimates for each of the health outcomes and behaviours. Each segment of the graph presents four different estimates for each outcome, health outcome or health behaviour, and the corresponding 95% confidence interval. These results show that the baseline estimates found using age left full time education were not just an artifact of the definition of education. I also estimate effects when education is defined as being in school after the age of 16, gaining A-levels or above and obtaining a degree.

Where the baseline estimates did not produce significant results neither are the estimates significant when the definition of education is changed. For example, for self-reported general health the estimate using age left full-time education is small and imprecisely estimated. When the other three definitions are used they remain insignificant. This is also the case for waist-to-hip ratio, hypertension, smoking and drinking behaviours.

The 2SLS estimates of education on the body size variables of BMI, overweight and obese remain significant when the definition is changed from the baseline of age left full-time education. The effects of being in education after the age of 16, gaining A-levels or above and obtaining a degree all result in higher point estimates reflecting the different education margin. This is what would be expected as they reflect being in education for more years, whereas the estimates in the first bar are the effect of one additional year. Typically, the point estimates of having an A-level or above and having a degree are larger than using being in education post–16 however they are not significantly different.

E. Mechanisms

To summarise the results so far I find significant 2SLS estimates for education on BMI for both men and women, and this results, for men, in a reduction in the probability in being obese and waist size. Cutler and Lleras-Muney (2006) pose a number of mechanisms that could explain the relationship between education and health. These include income and access to health care, improvements in labour market status, value of the future, information and cognitive skills, preferences, rank and social networks. Devereux and Fan (2011) directly examine the effect of education on income from this reform and Machin et al. (2012) investigate wages as a mechanism in the education and crime relationship. I am able to examine economic activity and social class which could be relevant for rank being a mechanism as mentioned in Cutler and Lleras-Muney (2006). I am not able to, however, distinguish between those mechanisms nor fully rule out that others do not matter but I do show some mechanisms that are important.

Table 6 shows the mechanisms considered here. Tables A6 and A7 show the effects for males and females respectively. Two measures of social class are presented in Column (1) and (2), these are a dummy variable for being in the professional class, and one for being in either the professional or managerial class. The second mechanism tested is economic activity. Column (3) presents estimates on being employed, column (4) on the probability of being unemployed, and in column (5) whether the respondent is inactive. In addition, Column (6) examines the role that depression might play. It is well known that depression might contribute to being overweight, Markovitz, Friedman, and Arent (2008). Therefore, I use whether someone has scored above four on the GHQ–12 survey. This gives an indication of a possible mental health issue[8]. Due to the results found on body size, columns (7) and (8) examine both height (cm) and weight (kg) to see which of those was driving the result.

Table 6 shows the reduced form, OLS and 2SLS estimates for all individuals. There are significant increases in being in either the professional or managerial social class. When the managerial class is not included, column (1), the 2SLS becomes insignficant. There is also an improvement in economic activity. The reform lead to an increase in the probability of employment and decreased the probability in unemployment as reflected in the reduced form estimates, the p-value of the test of the joint probabilty of the expansion and post-expansion cohorts being significant albeit only at the 10% level for being employed. The OLS estimates of education on economic activity are significant and the expected signs. When the labour market mechanisms are examined separately for each gender there a similar pattern emerges, however, the estimates for the sub-samples are not precisely estimated. I find no evidence of education affecting mental health for either men or women, this is in contract to Brunello et al. (2013) who did find.

Turning to height and weight, in the pooled sample (Table 6) it can be seen that weight is the primary driver of the reductions in body size. This is reflected quite strongly in both the reduced form and 2SLS estimates. The results for weight are less precise for men, although of a similiar magnitude. There appears to be a positive increase in height as a result for men. For women, there is no such effect on height and the weight estimates are of a similar magnitude, around 3.4 kg, as the pooled sample.

Although peer effects can not be directly tested, the HSE does contain information at the household level. Therefore, I am able to examine whether the education expansion had an effect on the partners of those as those who were affected by the reform. Table A8 presents the results of this analysis. There is no evidence that the education expansion

[8]For more background on the GHQ-12 used in the Health Survey of England please see: `http://www.hscic.gov.uk/catalogue/PUB13218/HSE2012-Ch4-Gen-health.pdf`

had an impact on the health of the partners who were affected. I restrict the sample to those who are either married or cohabiting. Panel A presents the estimates for BMI. I first show the OLS estimates for the effect of the individual's education and then for that of the partner. I find a significant and negative effect of both education and the partner's education on BMI across a range of different specifications. Column 1 presents the baseline estimates that restricts the age of the sample to those aged between 23 and 34. In column 2 this is relaxed and allow the individuals to be any age, this is to allow partners to have different ages, and to increase the sample size in order to attempt to gain more precision of the estimates. In column 3 individuals who are part of the expansion cohorts are excluded in order that any effect is driven by the partner and is not a combination of the individual and partner. Column 4 again excludes the individuals from being part of the expansion cohort and only includes those with a partner to be born after 1972. The partner's education is instrumented by the expansion cohorts of the partner and a post expansion cohort dummy. To some extent this is problematic as only the sample that does not have any sample restriction imposed produces a first stage F-statistic that is above 10 for both the BMI and variants thereof and for the waist measurement. However, none of the 2SLS estimates are significant for BMI, overweight, obese or waist circumference. Therefore, I am not able to detect the presence of peer effects via the partner.

F. Robustness Checks & Additional Results

I present robustness checks for the main significant results i.e. for the following dependent variables: BMI, BMI>30, and waist size in figures A2-A5. The figures show the estimate for a different specification with the corresponding 95% confidence interval. Column 1 presents the baseline estimates for comparison. The first set of changes are designed to see how sensitive the estimates are to changing the structure of the instruments.

Column 2 excludes the post education expansion dummy as an additional instrument this forces the identification to come from the cohorts that were effected just during the expansion years. Columns 3 and 4 changes the period of education expansion by including earlier cohorts as instruments. Column 5 reduces the number of instruments by combining the expansion cohorts into one (a single dummy for the 1972-1975 cohorts) and column 6 into two dummies (a dummy covering the 1972-1973 cohort and one for 1974-1975 cohort). Columns 7 and 8 revert back to the original instrument set and examine changing the specification of the age variables. In particular column 7 removes age completely, and column 8 replaces the polynomials of age with a set of age dummies. Specifications 9, 10 and 11 examine the effect of changing the regression specification with respect to age and year of birth. Column 9 presents the 2SLS estimates additionally including a cubic for year of birth. Column 10 additionally includes a quartic in age to the baseline specification, bar 11 presents the estimates when year of birth and age quartic are included.

One concern could be that age effects on health may have changed over cohorts, and controlling for age may not be sufficient to control for such effects. In order to address this specifications 12 includes an age cohort interaction as an additional control. In column 13 I additionally include the square of this interaction and in 14 also include a cubic, and in 15 a quartic. Another concern could be that the sample of individuals are too young to find effects, given some socioeconomic gaps in some health outcomes and behaviours can form very early on in life this is maybe unlikely. However, in order to examine whether changing the age of the sample matters in column 16 the sample restricted to those aged 24 to 33. The sample restricted to those aged 25 to 32 in column 17. The cohort window is examined in the next five columns. The sample is restricted to include individuals born between; 1965–1980 (column 18), 1967–1980 (column 19), 1967–1979 (column 20), 1967–1978 (column 21), 1968–1978 (column 22). Column 23 adds in a set of regional dummies. The health benefits of education could be enhanced if education results in individuals living in a better, more healthy area. Therefore, regional controls are added to take into account these effects.

In column 24 estimates are presented where LIML is used rather than 2SLS. This is because in over-identified models with weak instruments 2SLS will be biased and this is not the case with LIML. Therefore, a comparison between the 2SLS estimates and those estimated using LIML will also serve as a check on weak instruments. The next two bars on the figures are two falsification tests[9]are carried out. Here I replace the education expansion cohorts with a set of cohorts before the expansion. The final column examines the impact on the 2SLS when two additional controls are included. First, heath spending as a proportion of GDP when the respondent was aged 16 and second, unemployment rate at the time of the survey are included as additional controls. These are to caputure not only the potential health effect of the additional spending but also general sentiment of the government that would be with regards to promoting health as well as any effects on health that may results due to changes in the business cycle.

Taking the four figures together I underline two results. First, the baseline estimates are within the range of estimates that are presented using the alternative specifications and the estimates are robust to the inclusion of cohort age interactions. Second, when the expansion cohorts are combined into a single dummy the estimates for overweight, obese and waist are imprecisely estimated. However, this is not unexpected as this reduces the amount of variation that is being used. Similarly, while typically remaining significant, the estimates become less precise when the sample is made smaller when restrictions on age and cohorts are put in place. Third, the falsification tests are typically insignificant and much lower in magnitude that the baseline estimates.

[9]Two additional falsification tests are implemented. I run the baseline analysis using asthma and then skin conditions as the dependent variable. These are unlikely to be directly affected by education and therefore serve as a natural falsification test. Neither of the 2SLS estimates are significant. The coefficient on asthma is $-.0180$ with a standard error of 0.177. For skin conditions the 2SLS estimate is $-.004$ with a standard error of .010.

6. Conclusion

The late 1980s and early 1990s experienced a truly substantial increase in the participation of education beyond the compulsory school leaving age with the proportion of 18 year olds in full time education rising from around 17% in 1985 to over 35% in the late 1990s. This paper examines changes in health and health behaviour of cohorts that were affected by a large scale reform to post-compulsory schooling. This is in contrast with the vast majority of the literature that exploits changes in compulsory schooling laws. This reform had an impact across the entire education distribution, crucially affecting those at the bottom, and the top of the education distribution - increasing the years of completed schooling, the proportion with high school qualifications and those obtaining a college degree. An additional benefit in examining this particular reform was that it was recent, therefore, took place when many of the dangers regarding smoking, drinking and diet where all well-known, and in a more similar environment in terms of how much of a problem "bad" health behaviours were.

There is little evidence of an impact of education on self–reported measures of general health, having a long standing illness. Neither do I find any impact on some self-reported health behaviours - smoking and drinking behaviour. However, I find an effect of education in reducing body size. These results differ from much of the literature with very few papers finding an impact of education on overweight and obesity; exceptions include (Grabner, 2009; Webbink et al. 2010). The mechanisms that could account for these results are examined. Improvements in social class, and employment status are the two main drivers of the health improvements.

There are a number of possible reasons for the differences found here compared to the rest of the literature. First, given the reform is quite recent where obesity was becoming more prevalent, there is the potential for education to have an impact earlier on in life. Although the time of survey is similar to Clark and Royer (2013) who use the health survey of England 1991-2004, the education reforms would have had an effect when information about diet and the dangers of overweight were less well known. Second, the 2SLS estimates are picking up a broader group, due to the reform having effects at both ends of the distribution, than those studies that use either compulsory schooling law changes or college openings. The estimates of a reduction in the probability of being obese of around 7 percentage points as a result of an additional year of education are at the higher end of the effects that have been found, however, this study uses quite recent data where obesity has become a larger problem. While it is not possible to test the effect of peers directly, a further speculation for the cause these results could be that the peer group changed dramatically. The results testing different mechanisms suggest this could be a factor. Improvements in social class are found, which would have resulted in a different peer group.

This paper adds to the literature of this reform which has been used to show the impact of education on wages and crime, and now on improvements in health (BMI and a healthy diet). These results are notable in that the effects are shown for relatively young people at the time of survey, although they are not out of line with other studies. Furthermore, second only to smoking, obesity is the leading behavioural cause of death, and this paper has shown evidence that further improvements in education can go some way to prevent that.

References

Albouy, Valerie, and Laurent Lequien. 2009. "Does compulsory education lower mortality?," *Journal of Health Economics*, 28(1),155–168.

Arendt, Jacob N. 2009. "Does education cause better health? A panel data analysis using school reforms for identification," *Economics of Education Review*, 24(2): 149–160.

Blanden, Jo, and Stephen Machin. 2004. "Educational inequality and the expansion of UK higher education." *Scottish Journal of Political Economy* 51.2, 230-249.

Blanden, Jo, Gregg, Paul, and Stephen Machin. 2003. "Changes in educational inequality." Leverhulme Centre for Market and Public Organisation, University of Bristol, Department of Economics

Braakmann, Nils. 2011. "The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England." *Journal of Health Economics*, 30.4, 753-763.

Brunello, Giorgio, Fabbri, Daniele, and Margherita Fort. 2013. "The causal effect of education on body mass: Evidence from Europe." *Journal of Labor Economics*, 31.1, 195-223.

Clark, Damon, and Heather Royer. 2013. "The effect of education on adult mortality and health: Evidence from Britain." *The American Economic Review*, 103.6, 2087-2120.

Devereux, Paul J., and Wen Fan. 2011. "Earnings returns to the British education expansion." *Economics of Education Review*, 30.6, 1153-1166.

Etile, Fabrice, and Andrew M. Jones. 2011. "Schooling and smoking among the baby boomersan evaluation of the impact of educational expansion in France." *Journal of Health Economics*, 30.4, 811-831.

Gelman, Andrew, and Jennifer Hill. 2007. "Data Analysis using regression and multi-level/hieracrchical models." New York: Cambridge University Press

Gelman, Andrew, and Guido Imbens. 2014. "Why high-order polynomials should not be used in regression discontinuity designs." No. w20405. National Bureau of Economic Research.

Grabner, Michael. 2009. "The causal effect of education on obesity: Evidence from compulsory schooling laws." Available at SSRN 1505075

Gray, John, Jesson, David, and Mark Tranmer. 1993. "England and Wales Youth Cohort Study: boosting post-16 participation in full-time education: a study of some key factors." Employment Department ED Series Youth Cohort Report No. 20

Jrges, Hendrik, Reinhold, Steffen, and Martin Salm. 2011. "Does schooling affect health behavior? Evidence from the educational expansion in Western Germany." *Economics of Education Review*, 30.5, 862-872.

Kogan, Maurice, and Stephen Hanney. 2000. Reforming higher education. Vol. 50. Jessica

Kingsley Publishers.

Lleras-Muney, Adriana. 2005. "The relationship between education and adult mortality in the United States." *The Review of Economic Studies*, 72.1, 189-221.

Lochner, Lance. 2011. "Non-production benefits of education: Crime, health, and good citizenship." No. w16722. National Bureau of Economic Research.

Mazumder, Bhashkar. 2008. "Does education improve health? A reexamination of the evidence from compulsory schooling laws." *Economic Perspectives*, 32.2

Markowitz, Sarah, Friedman, Michael A., and Shawn M. Arent. 2008. "Understanding the relation between obesity and depression: causal mechanisms and implications for treatment." *Clinical Psychology: Science and Practice*, 15.1, 1-20.

Machin, Stephen and Marie, Olivier, and Vujić, Sunčica. 2012 "Youth crime and education expansion." *German Economic Review* 13.4, 366-384.

Oreopoulos, Philip. 2007. "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling." *Journal of public Economics* 91.11, 2213-2229.

Silles, Mary A. 2009. "The causal effect of education on health: Evidence from the United Kingdom." *Economics of Education Review* 28.1 (2009): 122-128.

Van Kippersluis, Hans, ODonnell, Owen, and Eddy van Doorslaer. 2011. "Long-run returns to education does schooling lead to an extended old age?" *Journal of Human Resources* 46.4, 695-721.

Walker, Ian, and Yu Zhu. 2008 "The College Wage Premium and the Expansion of Higher Education in the UK." *The Scandinavian Journal of Economics* 110.4, 695-709.

Webbink, Dinand, Martin, Nicholas G., and Peter M. Visscher. 2010. "Does education reduce the probability of being overweight?" *Journal of Health Economics*, 29.1, 29-38.

Figures and Tables

Figure 1: Proportion of 18 year olds in Full Time Education 1985-2000.
Source: Department of Education, Author's calculation.

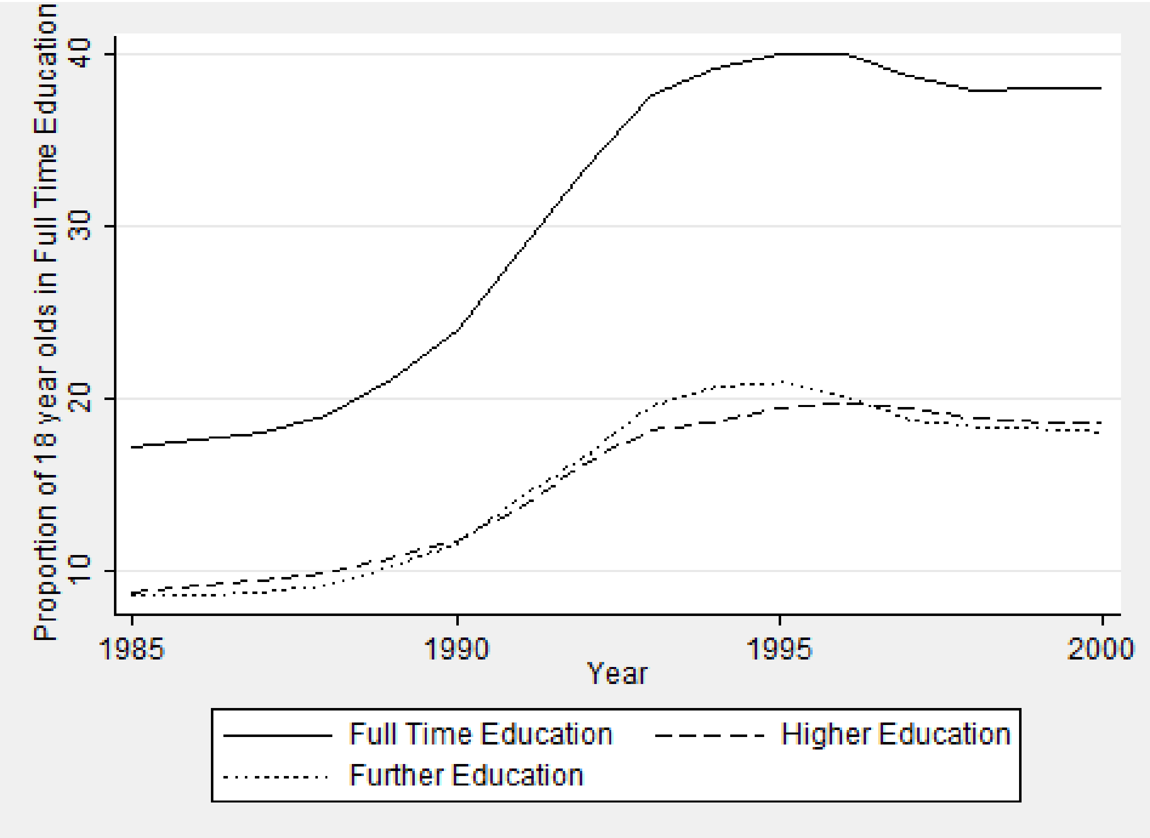
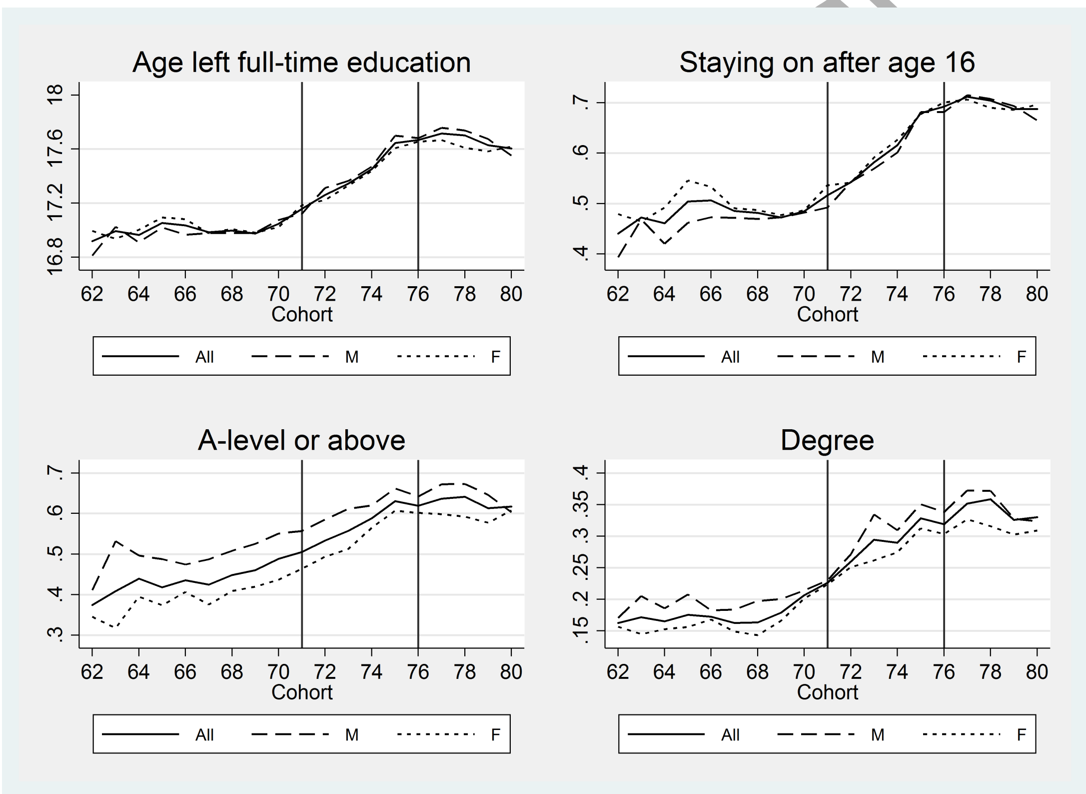


Figure 2: Education variables by birth cohort



source: Health survey of England

Figure 3: Self-reported health measures controlling for a cubic age profile

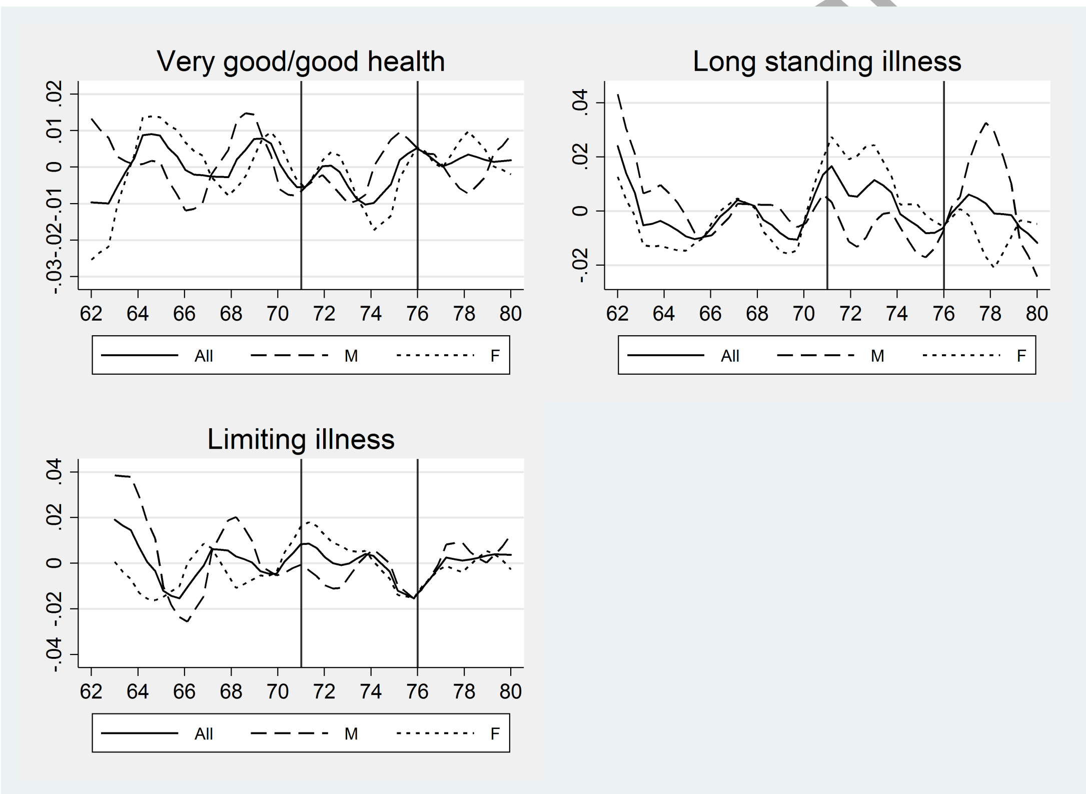


Figure 4: Body size controlling for a cubic age profile

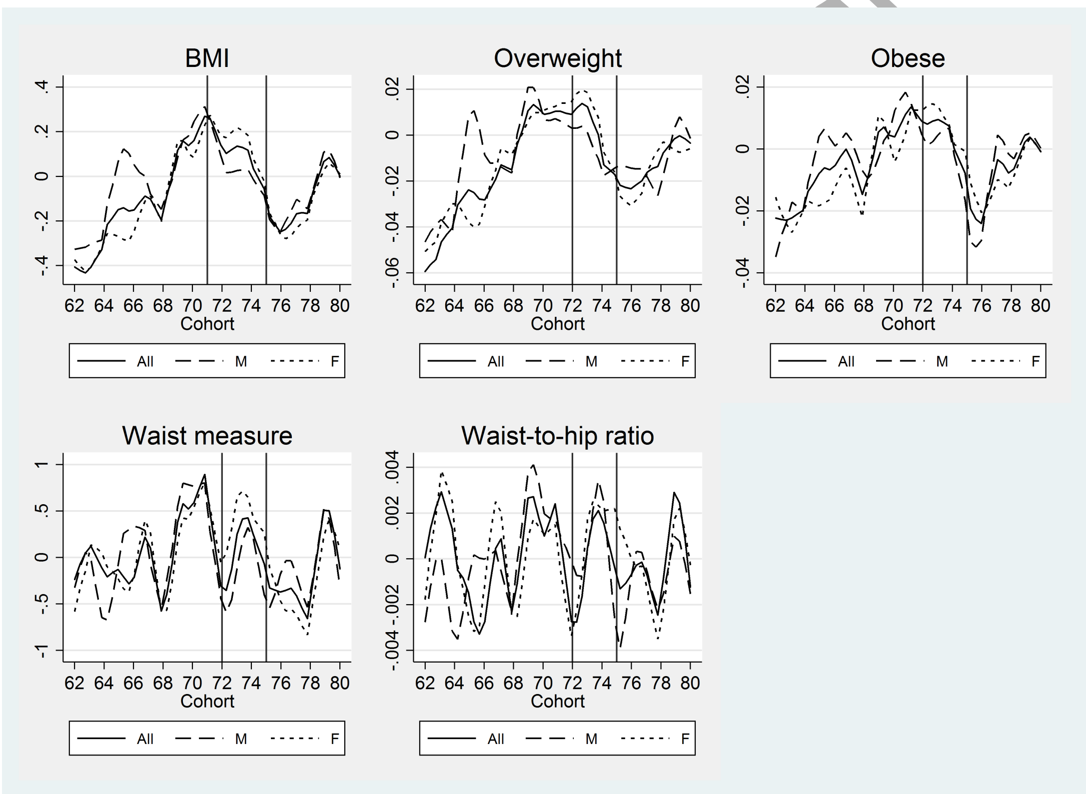


Figure 5: Blood Pressure controlling for a cubic age profile

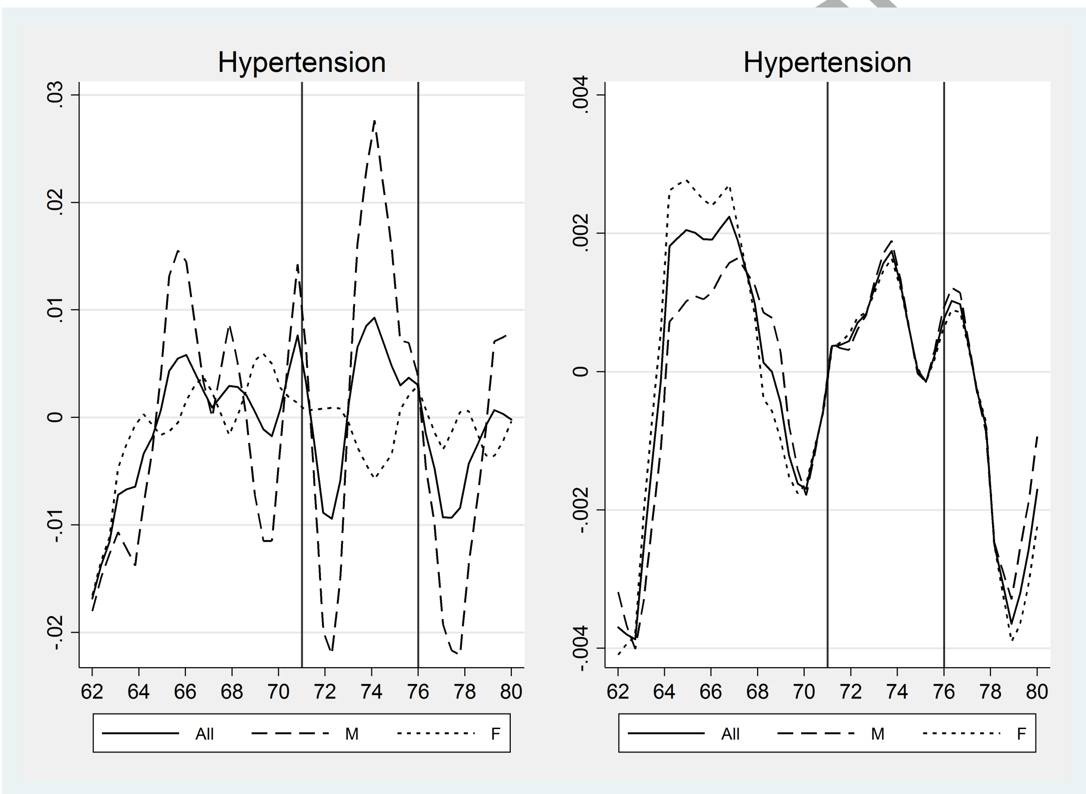


Figure 6: Health Behaviours controlling for a cubic age profile

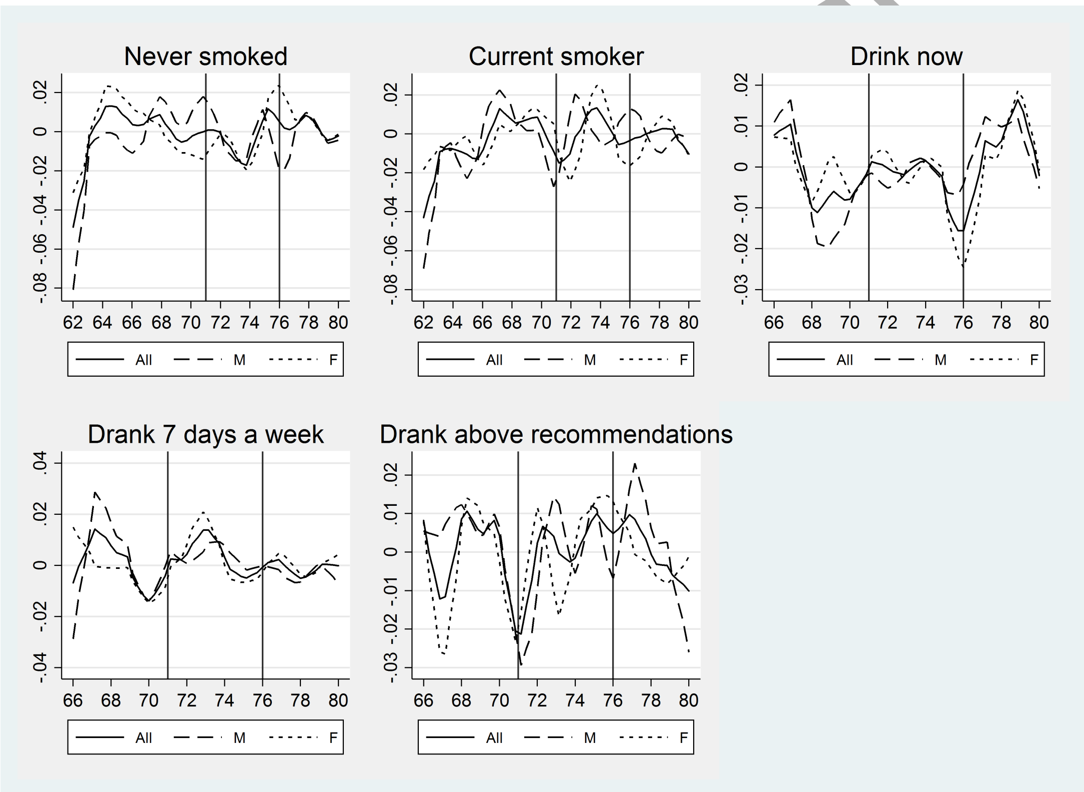
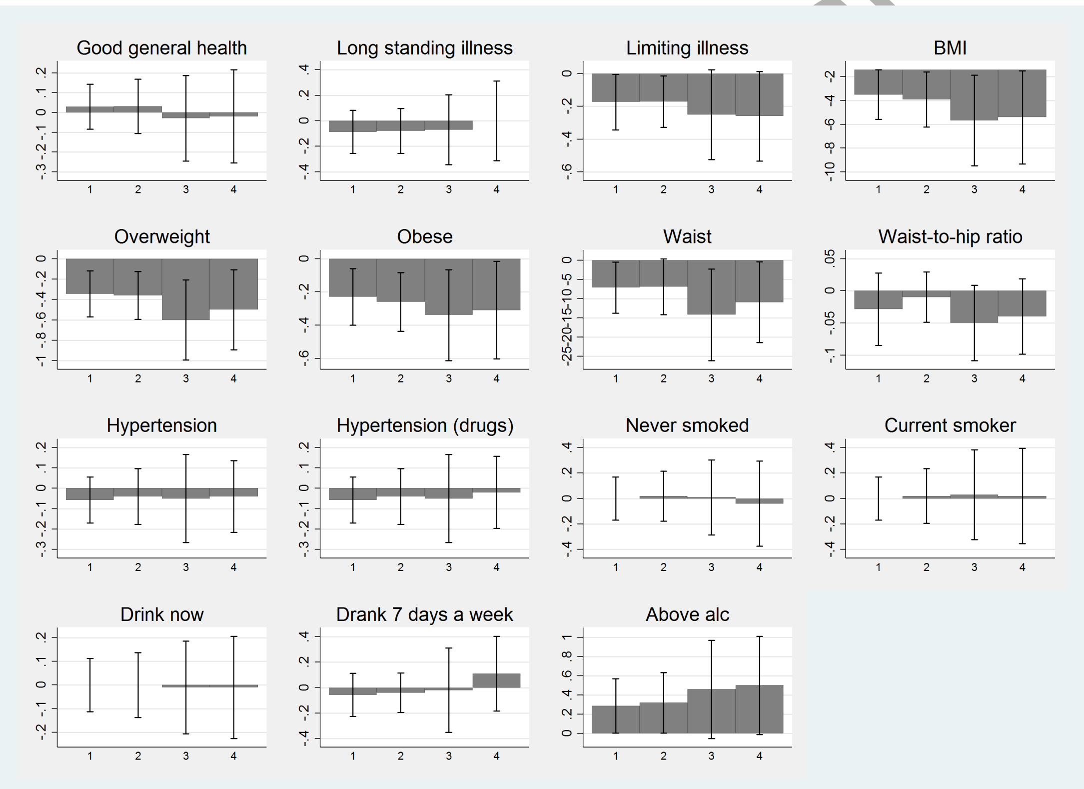


Figure 7: 2SLS estimates using alternative measures of education



Notes: Each bar corresponds to a 2SLS estimate and the wiskers represent the 95% confidence interval. In each segment each bar uses a different measure of education. Bar 1 uses the baseline measure, age left full time education. In order to be comparable to the other definitions of education the regression coefficient (and its corresponding standard errors) in bar 1 are multiplied by 2 standard deviations of the education variable, as suggested in Gelman and Hill (2007). In bar 2 education is defined as being "Post 16" i.e. in education after the age of 16. Bar 3uses an indicator for havig A-levels or above and bar 4 uses having a first degree or above.

Table 1: First stage results: The effect of education expansion on education achievement

	Age Left FTE	Post 16	A-Level +	Degree
Cohort 72	0.172***	0.0394***	0.0415***	0.0436***
	(0.0405)	(0.0139)	(0.0137)	(0.0118)
Cohort 73	0.258***	0.0806***	0.0568***	0.0687***
	(0.0450)	(0.0152)	(0.0151)	(0.0134)
Cohort 74	0.347***	0.109***	0.0719***	0.0551***
	(0.0478)	(0.0161)	(0.0161)	(0.0144)
Cohort 75	0.534***	0.170***	0.111***	0.0898***
	(0.0534)	(0.0178)	(0.0180)	(0.0166)
Post Expansion	0.546***	0.184***	0.0941***	0.0792***
	(0.0682)	(0.0226)	(0.0231)	(0.0215)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	21.11	19.77	8.302	8.475
P-value	0.00	0.00	0.00	0.00
Observations	30,339	30,339	31,504	31,504

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980. The F-stat is a test for the joint siginificance of the 1972 to 1975 and post expansion cohort dummy, the p-value corresponds to this F-test. The dependent variable in column (1) is is a variable defining the age the individual left full time education, column (2) is a dummy equal to 1 if the individual left school after age 16, column (3) is a dummy is the highest qualification achieved is an A-level (or equivalent) or above, and column (4) is a dummy equal to 1 representing whether the higest qualification of the individual is a degree.

Table 2: Self-Reported Health Outcomes

	Very Good/Good General Health	Long Standing Illness	Limiting Illness
Reduced Form			
Cohort 72	0.008	-0.005	-0.008
	(0.010)	(0.013)	(0.011)
Cohort 73	0.001	0.009	-0.019
	(0.011)	(0.014)	(0.012)
Cohort 74	-0.003	-0.009	-0.010
	(0.011)	(0.015)	(0.013)
Cohort 75	0.007	-0.018	-0.032**
	(0.013)	(0.017)	(0.014)
Post Expansion	0.022	-0.017	-0.039**
	(0.017)	(0.022)	(0.017)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	0.864	0.614	1.572
P-value	0.504	0.689	0.164
OLS	0.033***	-0.019***	-0.022***
	(0.001)	(0.002)	(0.002)
2SLS	0.005	-0.027	-0.057**
	(0.024)	(0.031)	(0.026)
Mean of Dep. Var (pre-expansion)	0.863	0.291	0.141
Observations	30,332	29,967	22,876
First Stage F stat	21.23	20.80	16.89
Hansen J (p-value)	0.385	0.695	0.638
Hausman Test (p-value)	0.360	0.745	0.141

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980. The F-test tests whether the coefficients on the excluded instruments are jointly equal to zero.

Table 3: Body Size

	BMI	Overweight	Obese	Waist	Waist to Hip Ratio
Reduced Form					
Cohort 72	-0.413**	-0.023	-0.010	-3.236***	-0.016***
	(0.198)	(0.021)	(0.016)	(0.745)	(0.005)
Cohort 73	-0.434*	-0.021	-0.012	-3.01***	-0.016***
	(0.254)	(0.027)	(0.020)	(0.947)	(0.006)
Cohort 74	-0.624**	-0.051*	-0.024	-3.072***	-0.013**
	(0.264)	(0.028)	(0.021)	(0.990)	(0.006)
Cohort 75	-0.793***	-0.060***	-0.047***	-2.936***	-0.013***
	(0.226)	(0.023)	(0.017)	(0.809)	(0.005)
Post Expansion	-0.954***	-0.065**	-0.056***	-2.223**	-0.007
	(0.244)	(0.026)	(0.019)	(0.867)	(0.005)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	3.536	2.064	2.613	4.151	2.903
P-value	0.004	0.067	0.023	0.001	0.013
OLS	-0.289***	-0.027***	-0.020***	-0.684***	-0.005***
	(0.021)	(0.002)	(0.002)	(0.075)	(0.000)
2SLS	-1.482***	-0.118***	-0.100***	-2.684**	-0.008
	(0.394)	(0.040)	(0.030)	(1.230)	(0.007)
Mean of Dep. Var (pre-expansion)	25.6	0.481	0.147	85.1	0.825
Observations	25,888	25,888	25,888	15,153	15,153
First Stage F stat	16.50	16.50	16.50	11.52	11.52
Hansen J (p-value)	0.805	0.900	0.994	0.00349	0.00871
Hausman Test (p-value)	0.00127	0.0187	0.00380	0.132	0.810

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, a dummy for and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980. The dependent variable in Column (1) is Body Mass Index (BMI) (2) and (3) are dummies for having a BMI>25 BMI>30. The F-test tests whether the coefficients on the excluded instruments are jointly equal to zero.

Table 4: Blood Pressure and Hypertension

	(1) Hypertension	(2) Hypertension
Reduced Form		
Cohort 72	-0.026***	-0.025***
	(0.009)	(0.009)
Cohort 73	-0.009	-0.004
	(0.012)	(0.011)
Cohort 74	-0.003	-0.002
	(0.013)	(0.013)
Cohort 75	-0.008	-0.009
	(0.015)	(0.015)
Post Expansion	-0.016	-0.018
	(0.019)	(0.019)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	1.717	1.730
P-value	0.127	0.124
OLS	0.0001	0.001
	(0.001)	(0.001)
2SLS	-0.016	-0.016
	(0.023)	(0.023)
Mean of Dep. Var (pre-expansion)	0.063	0.063
Observations	17,994	17,804
First Stage F stat	15.39	14.89
Hansen J (p-value)	0.093	0.092
Hausman Test (p-value)	0.456	0.420

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980. The F-test tests whether the coefficients on the excluded instruments are jointly equal to zero.

Table 5: Health Behaviours: Smoking and Drinking

	Never Smoked	Current Smoker	Drink Now	Drank 7 days in past week	Above levels
Reduced Form					
Cohort 72	-0.005	-0.008	0.010	0.006	0.028
	(0.014)	(0.014)	(0.009)	(0.012)	(0.021)
Cohort 73	-0.017	0.005	0.000	0.024*	0.025
	(0.015)	(0.015)	(0.010)	(0.014)	(0.023)
Cohort 74	-0.023	0.016	0.004	0.009	0.024
	(0.016)	(0.016)	(0.011)	(0.014)	(0.026)
Cohort 75	0.014	-0.006	0.001	-0.001	0.047*
	(0.018)	(0.018)	(0.013)	(0.015)	(0.028)
Post Expansion	-0.001	0.012	0.002	0.001	0.056*
	(0.023)	(0.023)	(0.017)	(0.018)	(0.033)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	1.464	0.697	0.316	1.059	0.793
P-value	0.198	0.626	0.903	0.381	0.555
OLS	0.039***	-0.069***	0.017***	-0.006***	0.007**
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
2SLS	0.005	0.001	0.003	-0.011	0.082
	(0.033)	(0.034)	(0.023)	(0.032)	(0.056)
Mean of Dep. Var (pre-expansion)	0.362	0.404	0.881	0.095	0.479
Observations	30,319	27,927	30,339	14,130	15,928
First Stage F stat	21.05	19.02	21.08	7.374	8.295
Hansen J (p-value)	0.121	0.481	0.814	0.279	0.816
Hausman Test (p-value)	0.317	0.036	0.542	0.770	0.161

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

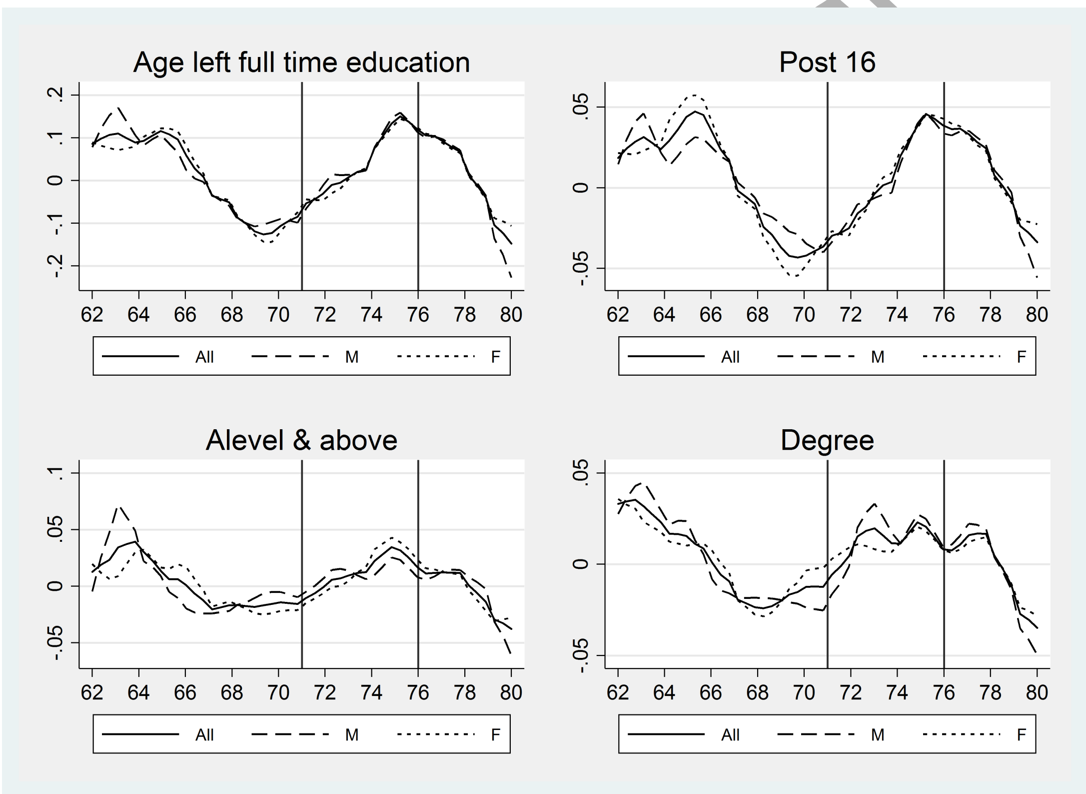
Table 6: Mechanisms

All	(1) Soc. Class Professional	(2) Soc. Class Professional/ Managerial	(3) Employed	(4) Unemployed	(5) Inactive	(6) GHQ 4+	(7) Height (cm)	(8) Weight (kg)
Reduced Form								
Cohort 72	0.001	0.001	-0.017	0.009	0.004	-0.013	-0.022	-0.741
	(0.006)	(0.013)	(0.012)	(0.007)	(0.011)	(0.010)	(0.273)	(0.459)
Cohort 73	0.017**	0.023	0.020	-0.011*	-0.010	-0.016	0.230	-0.341
	(0.008)	(0.015)	(0.013)	(0.006)	(0.012)	(0.012)	(0.302)	(0.520)
Cohort 74	0.005	0.029*	0.001	-0.012*	0.008	-0.010	-0.070	-1.186**
	(0.008)	(0.016)	(0.014)	(0.007)	(0.013)	(0.013)	(0.328)	(0.559)
Cohort 75	0.015	0.030	0.018	-0.001	-0.016	-0.016	0.086	-1.957***
	(0.009)	(0.018)	(0.016)	(0.008)	(0.014)	(0.014)	(0.373)	(0.629)
Post Expansion	0.021	0.028	0.001	-0.005	0.006	-0.049***	0.284	-2.537***
	(0.013)	(0.025)	(0.020)	(0.011)	(0.019)	(0.018)	(0.477)	(0.812)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	1.433	1.110	1.955	2.404	1.245	2.058	0.316	2.678
P-value	0.209	0.353	0.082	0.035	0.285	0.068	0.904	0.020
OLS	0.036***	0.137***	0.057***	-0.007***	-0.051***	-0.005***	0.611***	-0.272***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.040)	(0.068)
2SLS	0.026	0.064**	0.034	-0.011	-0.022	-0.024	0.123	-3.47***
	(0.017)	(0.031)	(0.028)	(0.015)	(0.026)	(0.026)	(0.706)	(1.208)
Mean of Dep. Var (pre-expansion)	0.053	0.324	0.748	0.054	0.192	0.157	169.1	73.5
Observations	28,757	28,757	30,317	30,317	30,317	27,621	28,791	26,990
First Stage F stat	20.16	20.16	21.06	21.06	21.06	18.59	18.17	18.04

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

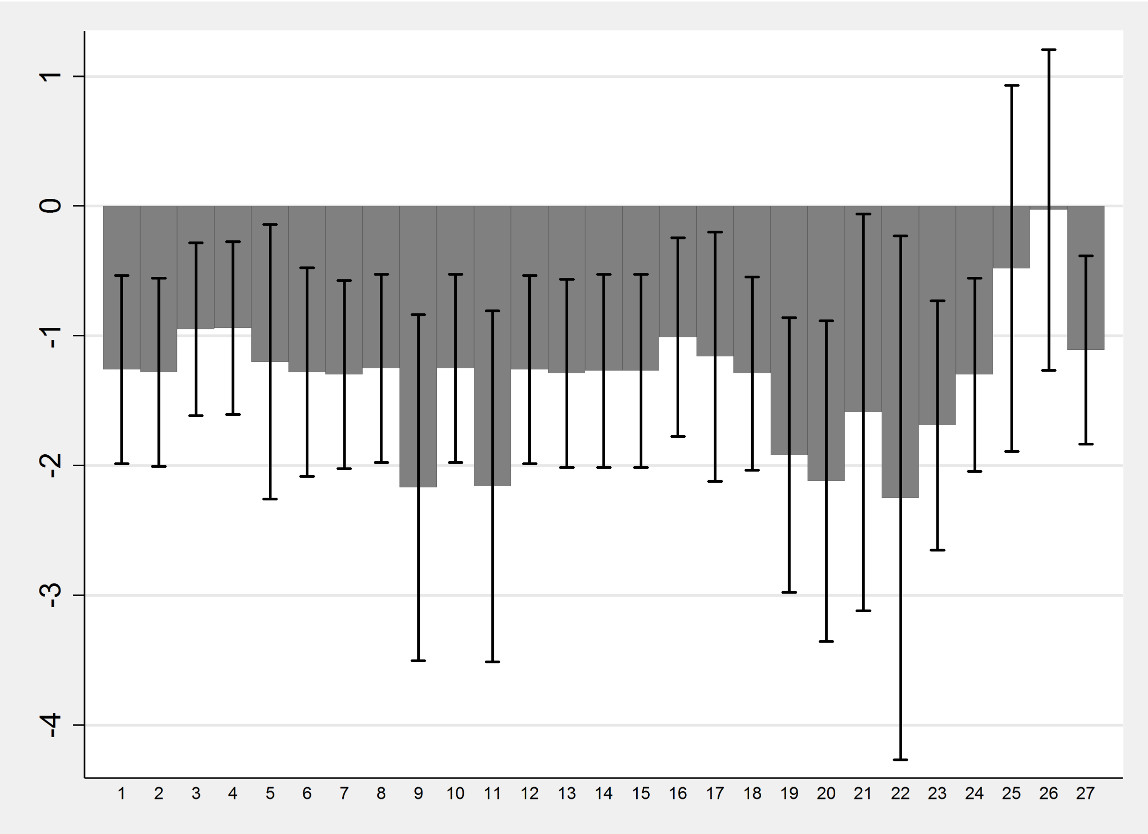
Appendix A

Figure A1: Education controlling for a cubic age profile



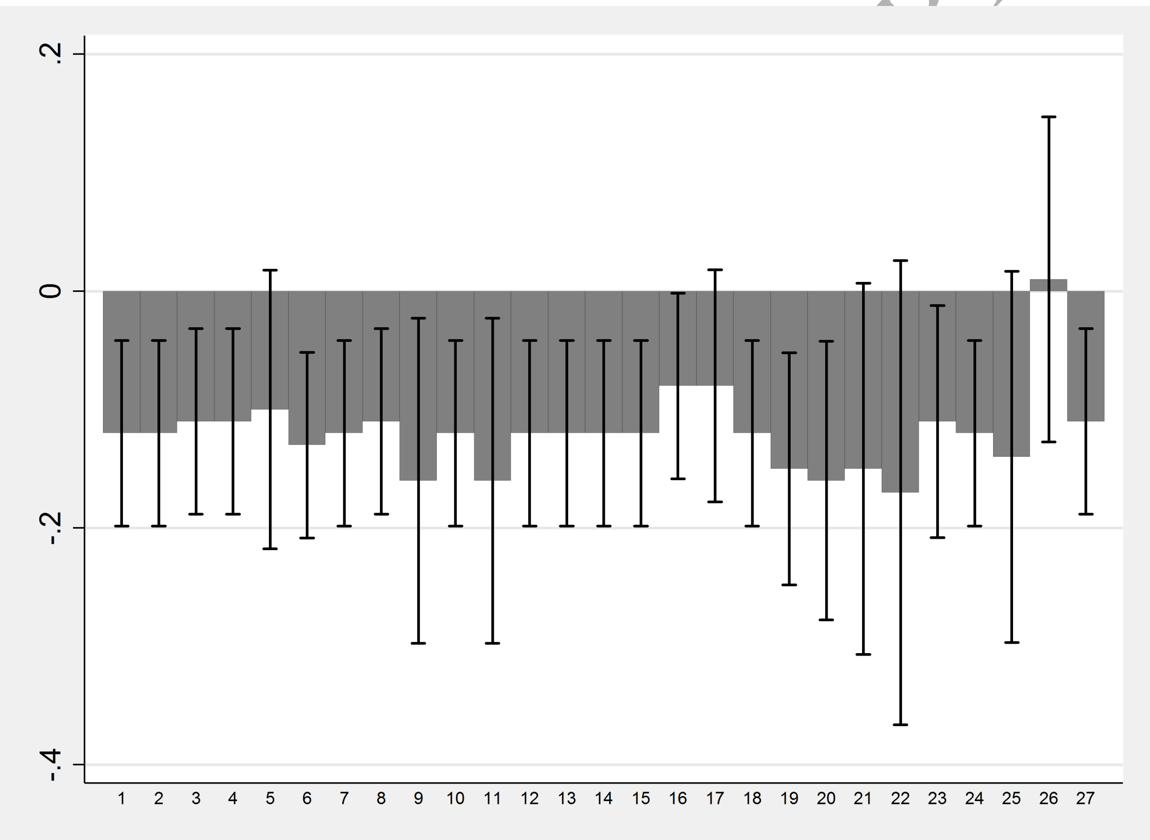
source: Health survey of England

Figure A2: 2SLS estimates of the effect of education on BMI using alternative specifications



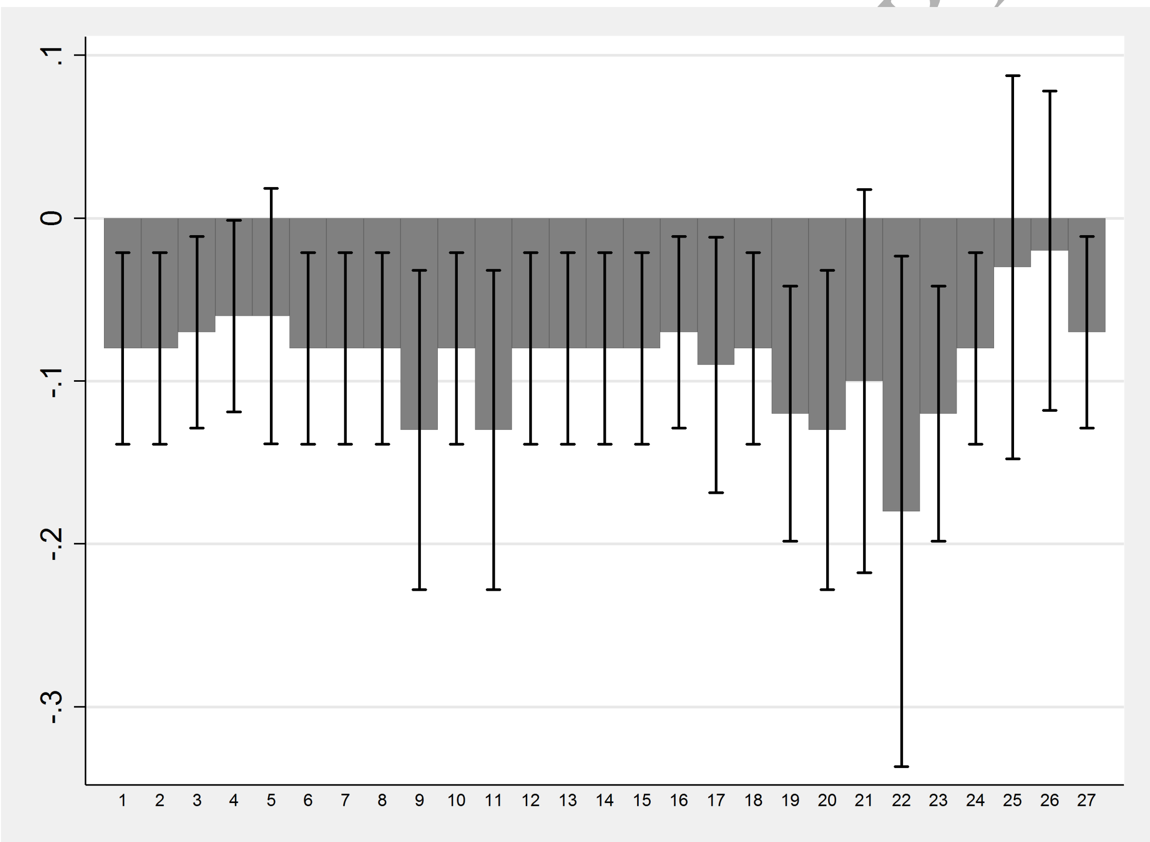
Notes: Each bar corresponds to a 2SLS estimate and the wiskers represent the 95% confidence interval. 1. Is the 2SLS baseline estimate. 2. Excludes post expansion dummy 3. Baseline + 1970 and 1971 cohort dummies 4. Baseline + 1971 cohort dummy. 5. Single 1972-75 chort dummy + post expansion dummy. 6. Cohort 1972-73 + Cohort 1974-75 + post expansion dummy. 7. No age variables are included. 8. Age dummies replace age specification in the baseline. 9. Year of birth cubic additionally included. 10. Age quartic additionally included to the baseline specification. 11. Year of birth cubic and age quartic both additionally included. 12. Cohort age interaction included. 13. Cohort age interaction and its square included. 14. Cohort age interaction, its square and cubic included. 15. Cohort age interaction, its square, cubic and quartic included. 16. The sample restricted to those aged 24 to 33. 17. The sample restricted to those aged 25 to 32. 18. The sample restricted to the cohorts 1965–1980. 19. The sample restricted to the cohorts 1966–1980. 20. The sample restricted to the cohorts 1967–1980. 21. The sample restricted to the cohorts 1967–1978. 22. The sample restricted to the cohorts 1968–1978. 23. Regional dummies included. 24. LIML is used instead of 2SLS. 25. The cohorts 1972-1975 are replaced by the cohorts 1966-1969. 26. The cohorts 1972-1975 are replaced by the cohorts 1962–1965. 27. Heath spending as a proportion of GDP when the respondent was aged 16 and unemployment rate at the time of the survey are included as additional controls.

Figure A3: 2SLS estimates of the effect of education on overweight using alternative specifications



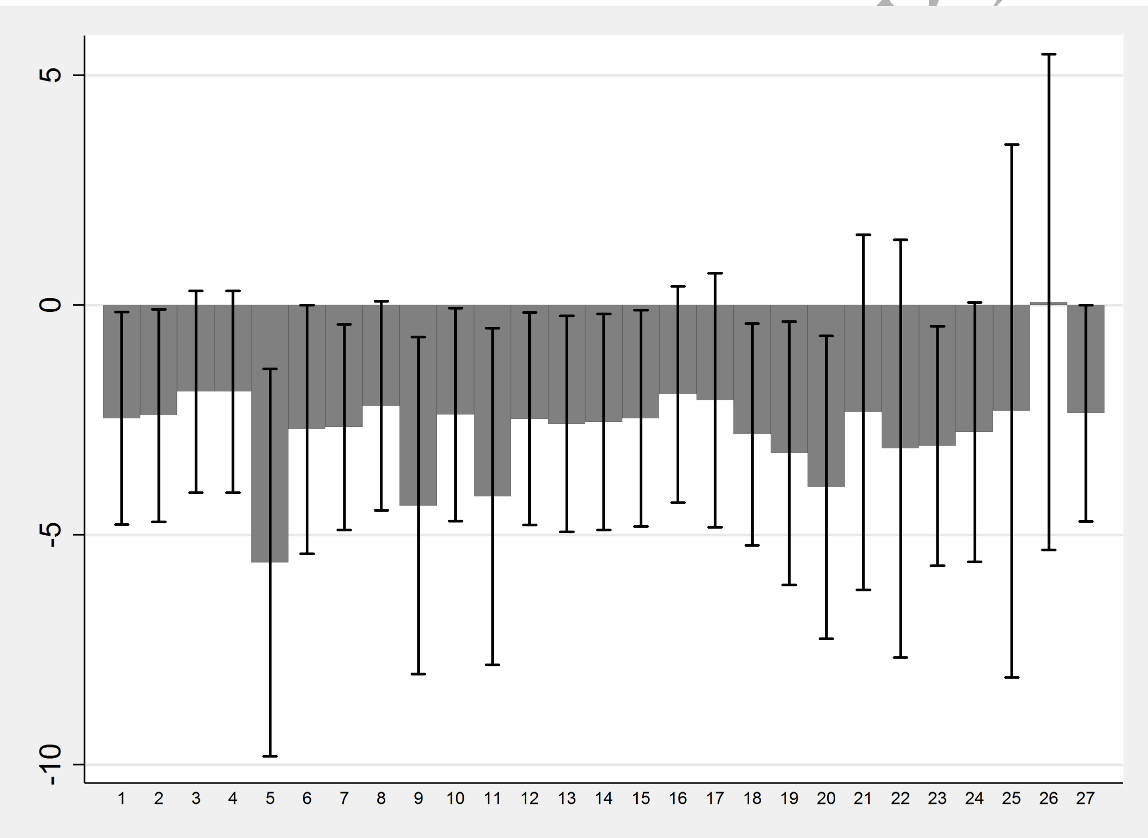
Notes: see notes to Figure A2.

Figure A4: 2SLS estimates of the effect of education on obese using alternative specifications



Notes: see notes to Figure A2.

Figure A5: 2SLS estimates of the effect of education on waist circumference using alternative specifications



Notes: see notes to Figure A2.

Table A1: Summary statistics of age and education variables by cohort

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cohort	Age	Max Age	Min Age	Age left full-time education	Post 16	*N* Columns (1)-(5)	A-level	Degree	*N* Columns (7)-(8)
1962	31.7	34	28	16.9	0.44	738	0.37	0.16	745
1963	31.1	34	28	17.0	0.47	938	0.41	0.17	952
1964	31.0	34	27	17.0	0.46	1224	0.44	0.17	1240
1965	30.7	34	26	17.1	0.50	1375	0.42	0.18	1396
1966	30.2	34	24	17.0	0.51	1549	0.44	0.17	1574
1967	29.9	34	24	17.0	0.49	1692	0.43	0.16	1746
1968	29.5	34	23	17.0	0.48	1936	0.45	0.17	1997
1969	29.3	34	23	17.0	0.47	2011	0.46	0.18	2083
1970	29.2	34	23	17.0	0.48	1861	0.49	0.21	1918
1971	29.3	34	23	17.2	0.52	1868	0.50	0.23	1957
1972	29.0	34	23	17.3	0.54	1936	0.54	0.26	2030
1973	28.8	34	23	17.3	0.58	1774	0.56	0.29	1858
1974	28.9	34	23	17.5	0.62	1876	0.59	0.29	1963
1975	28.5	34	23	17.6	0.68	1696	0.63	0.33	1767
1976	28.6	34	23	17.7	0.69	1597	0.62	0.32	1647
1977	28.3	34	23	17.7	0.71	1612	0.64	0.35	1686
1978	27.8	34	23	17.7	0.70	1791	0.65	0.36	1895
1979	27.3	33	23	17.6	0.69	1578	0.63	0.33	1699
1980	27.3	32	23	17.6	0.69	1287	0.62	0.34	1351
Total	29.1	34	24	17.3	0.57	30339	0.53	0.25	31504

Table A2A: Summary statistics of health variables by cohort

Cohort	Very good/ good health	*N*	Long standing illness	*N*	Limiting illness	*N*	BMI	*N*	Overweight	*N*	Obese	*N*	Waist (cm)	*N*	Waist-to-hip	*N*
1962	0.87	738	0.29	737	n/a		25.3	561	0.45	561	0.12	561	84.2	360	0.826	360
1963	0.86	937	0.26	937	0.18	89	25.2	727	0.45	727	0.12	727	84.3	397	0.828	397
1964	0.87	1223	0.27	1223	0.16	394	25.4	1005	0.46	1005	0.13	1005	84.1	591	0.823	591
1965	0.87	1375	0.26	1375	0.14	593	25.6	1112	0.48	1112	0.15	1112	84.1	621	0.820	621
1966	0.87	1549	0.26	1549	0.13	791	25.5	1252	0.46	1252	0.14	1252	83.4	660	0.814	660
1967	0.86	1691	0.28	1692	0.15	947	25.6	1419	0.47	1419	0.15	1419	84.7	736	0.823	736
1968	0.86	1936	0.27	1936	0.15	1226	25.4	1617	0.47	1617	0.13	1617	83.9	905	0.819	905
1969	0.87	2011	0.27	2011	0.14	1389	25.8	1746	0.51	1746	0.17	1746	86.2	1005	0.830	1005
1970	0.86	1861	0.27	1861	0.13	1421	25.6	1655	0.49	1655	0.15	1655	86.0	859	0.826	859
1971	0.85	1868	0.30	1867	0.15	1600	26.1	1633	0.51	1633	0.18	1633	87.9	801	0.835	801
1972	0.86	1936	0.27	1936	0.14	1701	25.7	1711	0.50	1711	0.17	1711	85.9	895	0.825	895
1973	0.85	1774	0.29	1774	0.13	1657	25.9	1559	0.50	1559	0.17	1559	86.8	911	0.829	911
1974	0.84	1875	0.28	1876	0.14	1875	25.9	1624	0.49	1624	0.17	1624	87.0	1030	0.831	1030
1975	0.85	1696	0.26	1695	0.12	1695	25.6	1504	0.47	1504	0.15	1504	86.2	969	0.825	969
1976	0.87	1595	0.26	1596	0.11	1596	25.4	1366	0.47	1366	0.14	1366	86.4	862	0.827	862
1977	0.86	1612	0.27	1612	0.13	1612	25.6	1398	0.48	1398	0.17	1398	86.4	932	0.828	932
1978	0.86	1790	0.26	1674	0.12	1674	25.6	1552	0.47	1552	0.16	1552	85.3	1024	0.820	1024
1979	0.85	1578	0.26	1461	0.13	1461	25.8	1376	0.47	1376	0.17	1376	86.6	901	0.825	901
1980	0.87	1287	0.24	1155	0.12	1155	25.5	1071	0.47	1071	0.16	1071	85.5	694	0.819	694
Total	0.86	30332	0.27	29967	0.13	22876	25.6	25888	0.48	25888	0.15	25888	85.7	15153	0.825	15153

Notes: Source: Health Survey of England 1991–2012, except for due to availability of the variable: limiting illness (1996–2011). The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A2B: Summary statistics of health variables by cohort

Cohort	Hypertension	N	Never smoked	N	Current smoker	N	Drink now	N	Drank 7 days of previous week	N	Drank above limits	N
1962	0.012	588	0.450	520	0.295	738	0.911	738	n/a		n/a	
1963	0.021	749	0.477	644	0.350	938	0.884	938	n/a		n/a	
1964	0.040	945	0.436	926	0.367	1224	0.882	1224	n/a		n/a	
1965	0.057	969	0.416	1099	0.372	1375	0.900	1375	n/a		n/a	
1966	0.066	1053	0.395	1276	0.363	1549	0.888	1549	0.088	441	0.486	251
1967	0.066	1124	0.424	1442	0.365	1691	0.889	1692	0.115	555	0.440	445
1968	0.076	1283	0.404	1687	0.370	1936	0.872	1936	0.105	725	0.494	666
1969	0.073	1311	0.399	1823	0.356	2009	0.875	2011	0.097	865	0.483	840
1970	0.075	1122	0.389	1710	0.361	1859	0.865	1861	0.074	900	0.504	936
1971	0.100	1065	0.348	1791	0.379	1867	0.870	1868	0.095	994	0.459	1067
1972	0.067	1108	0.354	1856	0.377	1936	0.863	1936	0.090	1077	0.492	1150
1973	0.090	995	0.361	1728	0.364	1773	0.853	1774	0.103	1041	0.496	1168
1974	0.100	1005	0.357	1876	0.364	1876	0.848	1876	0.082	1197	0.500	1336
1975	0.098	900	0.330	1696	0.406	1695	0.837	1696	0.069	1126	0.527	1274
1976	0.096	761	0.325	1596	0.400	1596	0.808	1597	0.070	1034	0.531	1329
1977	0.074	821	0.329	1609	0.395	1609	0.833	1612	0.069	1086	0.551	1404
1978	0.079	888	0.327	1790	0.415	1790	0.826	1791	0.057	1192	0.544	1558
1979	0.082	742	0.332	1575	0.390	1575	0.845	1578	0.063	1084	0.553	1394
1980	0.074	565	0.310	1283	0.408	1283	0.803	1287	0.059	813	0.550	1110
Total	0.072	17994	0.368	27927	0.376	30319	0.859	30339	0.080	14130	0.515	15928

Notes: Source: Health Survey of England 1991–2012, except for due to availability of the variable: drank 7 days of the previous week, and drank over limits on heaviest drinking day (1998–2012). The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A3: Cohort variation in pre-determined characteristics

	(1) Stillborn	(2) Infant Mortality Under 1 Year	(3) Infant Mortality Under 4 Weeks	(4) Birth's Outside Marriage
Cohort 1972	-0.0116	0.000247	0.000156	-0.278
	(0.0694)	(0.000304)	(0.00123)	(0.312)
Cohort 1973	0.0783	0.000478	0.000597	-0.269
	(0.109)	(0.000506)	(0.00226)	(0.481)
Cohort 1974	0.0881	0.000558	0.00161	-0.352
	(0.150)	(0.000593)	(0.00301)	(0.569)
Cohort 1975	0.0122	-1.91e-05	0.00189	-0.311
	(0.148)	(0.000525)	(0.00290)	(0.708)
Post EE Cohort	0.102	-0.00132	0.00382	-0.134
	(0.282)	(0.000883)	(0.00688)	(1.364)
Observations	23	23	23	23
R-squared	0.965	0.996	0.476	0.989
F-test	0.326	2.020	0.411	0.808
p-value	0.854	0.167	0.797	0.547

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies.

Table A4: 2SLS Estimates of The Effect of Education on Health by gender

	Self-Reported Health			Blood Pressure	
	Good General Health	Long Standing Illness	Limiting Illness	Hypertension	Hypertension + BP drug controls
Male	0.020	-0.047	-0.034	-0.049	-0.032
	(0.035)	(0.045)	(0.036)	(0.042)	(0.041)
Mean of Dep. Var (pre-expansion)	0.867	0.278	0.139	0.100	0.100
Observations	12,552	12,551	9,810	7.339	7.286
First Stage F stat	9.818	9.829	8.283	7.339	7.286
Female	-0.003	-0.028	-0.087**	-0.016	-0.025
	(0.034)	(0.043)	(0.039)	(0.023)	(0.023)
Mean of Dep. Var (pre-expansion)	0.860	0.268	0.143	0.037	0.0367
Observations	17,780	17,416	13,066	10,436	10,246
First Stage F stat	11.01	10.59	8.372	8.607	8.120
			Body Size		
	BMI	Overweight	Obese	Waist	Waist to Hip
Male	-1.447***	-0.127**	-0.131***	-3.739**	-0.017**
	(0.516)	(0.060)	(0.045)	(1.655)	(0.0083)
Mean of Dep. Var (pre-expansion)	25.9	0.555	0.141	91.5	0.882
Observations	11,579	11,579	11,579	6,319	6,319
First Stage F stat	7.376	7.376	7.376	5.390	5.390
Female	-1.075**	-0.0990**	-0.0497	-1.879	0.00425
	(0.522)	(0.0487)	(0.0366)	(1.417)	(0.00729)
Mean of Dep. Var (pre-expansion)	25.29	0.42	0.15	80.53	0.78
Observations	14,309	14,309	14,309	8,834	8,834
First Stage F stat	10.29	10.29	10.29	7.548	7.548

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A5: 2SLS Estimates of The Effect of Education on Health Behaviours by gender

	Never Smoked	Current Smoker	Drink Now	Drank 7 days	Above Recommend
Male	-0.062	0.060	0.015	-0.015	0.090
	(0.050)	(0.055)	(0.028)	(0.048)	(0.063)
Mean of Dep. Var (pre-expansion)	0.330	0.408	0.919	0.123	0.569
Observations	12,542	11,714	12,555	6,702	6,849
First Stage F stat	9.751	8.591	9.775	4.262	6.393
Female	0.048	-0.022	-0.015	-0.023	0.024
	(0.045)	(0.045)	(0.035)	(0.040)	(0.088)
Mean of Dep. Var (pre-expansion)	0.385	0.400	0.853	0.067	0.404
Observations	17,777	16,213	17,784	7,428	9,079
First Stage F stat	10.90	10.03	10.91	3.690	3.166

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A6: Mechanisms, male sample

All	(1) Soc. Class Professional	(2) Soc. Class Professional/ Managerial	(3) Employed	(4) Unemployed	(5) Inactive	(6) GHQ 4+	(7) Height (cm)	(8) Weight (kg)
Reduced Form								
Cohort 72	-0.002	0.004	-0.005	0.006	-0.002	0.001	0.309	-0.813
	(0.011)	(0.021)	(0.015)	(0.011)	(0.010)	(0.015)	(0.297)	(0.610)
Cohort 73	0.032**	0.020	0.013	-0.002	-0.009	0.002	0.038	-0.886
	(0.013)	(0.023)	(0.016)	(0.012)	(0.011)	(0.016)	(0.319)	(0.664)
Cohort 74	0.019	0.041	0.020	-0.0074	-0.012	-0.037**	0.374	-1.249*
	(0.014)	(0.025)	(0.017)	(0.012)	(0.012)	(0.017)	(0.363)	(0.735)
Cohort 75	0.022	0.037	0.020	-0.001	-0.016	-0.014	0.937**	-1.526*
	(0.016)	(0.028)	(0.019)	(0.014)	(0.014)	(0.020)	(0.406)	(0.835)
Post Expansion	0.037*	0.023	0.003	0.003	-0.002	-0.032	0.526	-2.538**
	(0.022)	(0.037)	(0.026)	(0.019)	(0.019)	(0.026)	(0.543)	(1.079)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	1.715	0.879	0.919	0.314	0.798	1.476	1.490	1.140
P-value	0.127	0.494	0.467	0.905	0.551	0.194	0.189	0.336
OLS	0.045***	0.141***	0.027***	-0.012***	-0.015***	-0.004*	0.539***	0.065
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.044)	(0.091)
2SLS	0.040	0.071	0.038	-0.005	-0.030	-0.033	1.558**	-3.096*
	(0.026)	(0.044)	(0.033)	(0.025)	(0.023)	(0.034)	(0.737)	(1.621)
Mean of Dep. Var (pre-expansion)	0.072	0.353	0.868	0.074	0.053	0.124	176.6	81.0
Observations	12,287	12,287	12,544	12,544	12,544	11,569	11,883	11,637

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A7: Mechanisms, female sample

All	(1) Soc. Class Professional	(2) Soc. Class Professional/ Managerial	(3) Employed	(4) Unemployed	(5) Inactive	(6) GHQ 4+	(7) Height (cm)	(8) Weight (kg)
Reduced Form								
Cohort 72	0.003	-0.002	-0.022	0.012	0.005	-0.024	-0.162	-0.630
	(0.008)	(0.017)	(0.018)	(0.008)	(0.017)	(0.014)	(0.237)	(0.563)
Cohort 73	0.003	0.024	0.024	-0.018***	-0.008	-0.028*	-0.105	-0.504
	(0.009)	(0.020)	(0.0191)	(0.006)	(0.019)	(0.016)	(0.279)	(0.672)
Cohort 74	-0.007	0.021	-0.007	-0.0174*	0.018	0.008	-0.272	-1.142
	(0.009)	(0.021)	(0.020)	(0.008)	(0.020)	(0.018)	(0.292)	(0.699)
Cohort 75	0.008	0.024	0.024	0.000	-0.024	-0.018	-0.313	-2.102***
	(0.011)	(0.024)	(0.023)	(0.010)	(0.022)	(0.020)	(0.335)	(0.779)
Post Expansion	0.006	0.031	-0.002	-0.009	0.012	-0.062**	-0.108	-2.964***
	(0.015)	(0.033)	(0.029)	(0.012)	(0.028)	(0.025)	(0.430)	(1.009)
F-Test of Joint Significance of Cohort 1972, 1973, 1974, 1975, & Post EE Dummy	0.717	0.496	1.595	4.114	1.227	3.221	0.415	2.098
P-value	0.611	0.779	0.158	0.001	0.293	0.007	0.839	0.063
OLS	0.030***	0.133***	0.078***	-0.003***	-0.076***	-0.006***	0.587***	-0.611***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.037)	(0.085)
2SLS	0.007	0.057	0.047	-0.018	-0.030	0.001	-0.652	-3.415**
	(0.021)	(0.044)	(0.043)	(0.019)	(0.041)	(0.039)	(0.677)	(1.456)
Mean of Dep. Var (pre-expansion)	0.038	0.303	0.662	0.0402	0.292	0.182	163.8	67.7
Observations	16,470	16,470	17,773	17,773	17,773	16,052	16,908	15,353
First Stage F stat	9.920	9.920	10.96	10.96	10.96	9.362	9.726	10.44

Notes: Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. The sample is aged between 23 and 34 and includes cohorts born between 1962 and 1980.

Table A8: Analysis of the effect of the education of a partner on selected health outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: BMI					
OLS					
Age left FT Education	-0.229***	-0.264***	-0.247***	-0.243***	-0.177***
	(0.0357)	(0.0259)	(0.0288)	(0.0313)	(0.0552)
Age partner left FT Education	-0.190***	-0.210***	-0.215***	-0.210***	-0.213***
	(0.0366)	(0.0264)	(0.0292)	(0.0318)	(0.0548)
2SLS					
Age partner left FT Education	-0.329	0.173	-0.00427	-0.0134	-0.835
	(0.566)	(0.488)	(0.798)	(0.635)	(1.546)
Panel B: Overweight					
OLS					
Age left FT Education	-0.029***	-0.026***	-0.025***	-0.025***	-0.021***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
Age partner left FT Education	-0.009**	-0.011***	-0.011***	-0.011***	-0.011**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
2SLS					
Age partner left FT Education	-0.078	-0.025	0.029	-0.054	-0.044
	(0.062)	(0.051)	(0.086)	(0.068)	(0.161)
Panel C: Obese					
OLS					
Age left FT Education	-0.015***	-0.019***	-0.019***	-0.019***	-0.015***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)
Age partner left FT Education	-0.014***	-0.015***	-0.015***	-0.015***	-0.018***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)
2SLS					
Age partner left FT Education	-0.070	-0.001	0.015	0.034	0.105
	(0.047)	(0.041)	(0.073)	(0.056)	(0.145)
Observations	11,319	22,238	14,410	17,753	5,191
First Stage F stat	8.847	11.76	3.619	6.475	1.267
Panel D: Waist					
OLS					
Age left FT Education	0.055	-0.110	-0.150	-0.062	0.304
	(0.277)	(0.212)	(0.259)	(0.237)	(0.465)
Age partner left FT Education	-0.102	-0.126	-0.174	-0.207	0.230
	(0.275)	(0.210)	(0.257)	(0.234)	(0.450)
2SLS					
Age partner left FT Education	3.404	2.294	5.123	0.760	-0.772
	(4.631)	(4.239)	(8.115)	(5.891)	(15.25)
Observations	9,837	19,722	12,575	15,681	4,753
First Stage F stat	8.234	10.34	2.606	5.329	0.936

Notes: the sample is restricted to those who are either married or cohabiting. Robust standard errors in parenthesis, *, ** and *** respectively denote significance at the 10, 5 or 1 percentage level. All specifications include a cubic polynomial in age, quadratic in year of birth, and year of survey dummies. 2SLS estimates also include a cubic polynomial in age of the partner, quadratic in year of birth of the partner.

Column (1) restricts the age of the sample to those aged between 23 and 34. 2SLS estimates instrument the education of the partner by the expansion cohorts and a post expansion cohort dummy of the partner. Column (2) imposes no age restriction. Column (3) includes those of any age and excludes individuals born from 1972-975. Column (4) includes those of any age and only those with a partner born after 1972. Excluded are individuals born from 1972–1975. Column (5) includes those of any age and only those with a partner born after 1972.